

An Investigation into the Impact of Website Design and Aesthetics on the Perception
of Bias in News Articles

A thesis Submitted to the
University of Dublin,
Trinity College
for the Degree of
Doctor in Philosophy



Brendan Spillane
Knowledge and Data Engineering Group,
School of Computer Science and Statistics, Trinity College Dublin

2020

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor Vincent Wade) Principle Co-Supervisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor Séamus Lawless) Assistant Co-Supervisor

Approved for the Trinity College Dublin Committee on Graduate Studies

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

_____ Brendan Spillane March 2019

Acknowledgments

There are too many reasons to be thankful to the people who contributed to and supported this endeavour, but fortunately not too many to thank.

First, I would like to thank my supervisors, Professor Vincent Wade and Professor Séamus Lawless, without whose support, encouragement, guidance and knowledge I could not have completed this work. I continued to enjoy their support throughout the ups and most importantly the downs. I am forever grateful for this and hope that this thesis and the other work we have committed to completing begins to repay the debt.

To my mother Margaret and my late father Michael, whose death during the course of this PhD, put into stark contrast the important things in life. Their love and support and the passion they instilled in each of us for the value of education can only be beaten by the sacrifices they made to avail us of it. To my brothers and sisters, blood will always run thicker than water.

If the value of a man's life is measured by his friendships, then I am rich. Each of you know who you are.

Brendan Spillane

Dedicated to the memory of my father Michael Spillane - 4th January 1947 to 31st July 2015,
and to my mother Margaret Spillane.

Also dedicated to Professor Séamus Lawless, my friend and the Co-Supervisor of this PhD, who died on
the descent after summiting Mt Everest in the area known as The Balcony on the morning of May 16th
2019.

Abstract

Most journalists and editors agree that it is virtually impossible to produce and disseminate news that is completely unbiased, and even if they did succeed, each consumer's personal biases would still influence the message. News, in one form or another, is biased. On that the research community agrees. What the research community has yet to agree, is a succinct yet encompassing definition, a categorisation of its various forms, their manifestation, and their impact, and how users arrive at or make judgements of bias. The subjective and constantly changing nature of the construct means that agreement on these, or even what exactly bias is, may never be forthcoming. Despite this, there has been a broad range of research on bias in the dissemination of news in traditional mediums. Yet, bias in the dissemination of news online has yet to receive the same level of attention. The most obvious lacuna in the literature is the impact of the visual presentation of news websites on perceived bias in the news articles they contain.

This thesis begins with a review of major authors in the domain lamenting the lack of serious underpinning theoretical research. It then provides a comprehensive literature review, including detailed classifications of biases that impact the production and dissemination of news. Due to a lack of theoretical and empirical models and frameworks explaining how judgements of bias are made or how it should be measured, the overarching domain of credibility is explored which has an abundance of models, theories and frameworks explaining how users form judgements of credibility online. These are also well supported with established underpinning theories. As no empirical frameworks were identified to aid in the design, development and deployment of experiments to measure credibility or any of its individual dimensions and measures such as bias, a first of its kind empirical framework and accompanying experiment platform to support such experiments was created.

Bias has been shown to be a core dimension and measure of credibility, especially when users are judging news on the WWW. Research has also shown that the visual presentation of websites affects judgements of credibility. This research was instigated based on the supposition that judgements of bias in news articles are similarly impacted by their visual presentation. The first experiment, described in this thesis, demonstrated that perceived bias in a news article could be impacted by the inclusion or omission of individual features. The second experiment investigated the design characteristic of professionalism. It found that as the level of professionalism in the aesthetics of a news webpage decreased, the perception of bias in the news article it contained increased, and vice versa. This effect was common across all webpage/article combinations tested and articles ranging from relatively unbiased to extremely biased. The final experiment found that based on this knowledge, it is possible to predictively impact the perception of bias by adapting another visual characteristic of a news webpage, visual quality. This demonstrates that perceived bias in a news article is at least partially judged heuristically in a similar fashion to credibility of which it is a core dimension and measure. This is significant for both news website designers and news consumers.

Table of Contents

Declaration.....	iii
Acknowledgments	iv
Abstract.....	vi
Table of Contents.....	vii
Table of Figures	xxii
Table of Tables	xxiv
Abbreviations.....	xxvii
Definitions of Key Terms	xxix
Website’s Design	xxix
Webpage’s Aesthetic	xxix
Characteristics of a Website’s Design	xxix
Features of a Website’s Design.....	xxix
Dimensions of Credibility.....	xxx
Measures of Credibility.....	xxx
Bias – A Core Dimension and Measure of Credibility	xxx
1 Introduction	1
1.1 Motivation	4
1.2 Research Question and Objectives	8
1.2.1 Research Question	8
1.2.2 Research Objectives.....	9
To address the research question outlined above, the following research objectives were devised for this thesis:.....	9
1.3 Research Contributions.....	10
1.3.1 Published Conference Papers.....	12
1.3.2 Journal Papers being Prepared for Submission.....	14
1.4 Research Methodology	15
1.5 Thesis Structure	16
1.5.1 Combined Experiments Two and Three	16
1.6 The Thesis Focus – Bias, as a Dimension and Measure of Credibility	16
1.7 Bias in the Research	19
1.8 Application Domain and Findings.....	19
2 Related Work – State of the Art.....	20
2.1 Introduction to State of the Art.....	20
2.2 Credibility	21
2.2.1 Introduction to Credibility	21
2.2.2 Origins of Research into Credibility	22

2.2.3	Credibility Research on the Internet	22
2.2.4	Definitions of Credibility	23
2.2.4.1	Perceptual Variable	23
2.2.4.2	Believability	23
2.2.4.3	Dimensions	24
2.2.4.4	Philosophical	24
2.2.5	Two Sides to the Coin: Defining the Construct of Credibility, and Measuring Credibility of Source, Medium, and Message.....	24
2.2.6	The Three Meanings of ‘Dimension of Credibility’	25
2.2.6.1	Bias as a Core Dimension of Credibility	27
2.2.6.2	Bias as a Core Measure of Credibility.....	28
2.2.7	Dual-Process Models Persuasion	28
2.2.7.1	Judgements of Information Via the Central Route, Systematic Strategy, and Controlled Information Process of the ELM, HSM, and C&AHIP Respectively	29
2.2.7.2	Judgements of Information Via the Peripheral Route, Heuristic Strategy, and Automatic Human Information Process of the ELM, HSM, and C&AHIP Respectively	29
2.2.7.3	Judgements of Credibility and Bias on the WWW	29
2.2.7.3.1	The Elaboration Likelihood Model (ELM).....	30
2.2.7.3.2	The Heuristic Systematic Model (HSM).....	32
2.2.7.3.3	Controlled and Automatic Human Information Processes (C&AHIP).....	34
2.2.8	Human-Centric Models, Theories, Frameworks, and Schematics (MTFS) for Judging or Measuring Credibility Online.....	35
2.2.9	Connecting the Underlying Theory of the ELM, HSM, and the C&AHIP, to the MTFS of Credibility Judgement, and the Visual Presentation of News Webpages	41
2.2.9.1	Connecting the MTFS of Credibility Judgement to Individual Features and Characteristics of a News Website’s Design, Reflected in a news Webpage’s Aesthetic	41
2.2.9.2	Examples of Individual Features and/or Characteristics by which a User May Make Judgements of Credibility via the Peripheral Route, Heuristic Strategy, or the Automatic Human Information Process.....	42
2.2.10	News Website Design, News Webpage Aesthetics, and the Production of News Online...43	
2.2.11	The Impact of Features and Characteristics of a Website’s Design, Reflected in a Webpage’s Aesthetic, on the Perception of Credibility	44
2.2.12	Motivating Factors Based on Issues with Current Approaches to Measuring Credibility Online, and any of its Dimensions such as Bias, and the Resulting Requirements for an Open Framework.....	49
2.2.12.1	Framework Requirements	49
2.2.12.2	Repository Requirements	50

2.2.12.3	Platform Requirements.....	51
2.2.13	Experimental Frameworks for Detecting Credibility or any of its Dimensions or Measures	53
2.2.14	Conclusions on Credibility	53
2.2.15	Summary on Credibility.....	54
2.3	Bias	54
2.3.1	Introduction to Bias.....	55
2.3.2	Bias, The Lack of Underlying Theoretical Research.....	55
2.3.3	Definitions of Bias	58
2.3.3.1	Difficulties in Measuring Bias	58
2.3.3.2	Bias Criteria: Volitional, Influential, Threatening, Choice, Sustained, Systematic and Effective	59
2.3.3.3	Definitions of Cognitive Bias.....	60
2.3.3.4	Non-Definitions of Bias	60
2.3.3.5	Domain Specific Definitions of Bias	61
2.3.4	Bias in the Three Main Stages of the News Cycle: Production, Dissemination, and Consumption	62
2.3.4.1	Bias in the Production of News.....	64
2.3.4.2	Bias in the Dissemination of News	65
2.3.4.3	Cognitive Bias in the Consumption of News.....	65
2.3.5	Bias in the Dissemination of News in Traditional Mediums: Print, Radio, and Television	65
2.3.5.1	Studies on Bias in Print Newspapers and News Magazines	66
2.3.5.2	Studies on Bias in Radio News	66
2.3.5.3	Studies on Bias in Television News.....	66
2.3.6	Bias in the Dissemination of News on the WWW.....	67
2.3.6.1.1	Headline Bias in News on the WWW.....	69
2.3.6.1.2	Coverage Bias in News on the WWW.....	70
2.3.6.1.3	Photograph Bias in News on the WWW.....	70
2.3.6.1.4	Presentation Bias in News on the WWW	70
2.3.6.1.5	Visual Bias in News on the WWW.....	71
2.3.6.1.6	Position Bias in News on the WWW	71
2.3.6.1.7	Design Bias in News on the WWW.....	72
2.3.7	Bias Frameworks: Theoretical	73
2.3.8	Bias Frameworks: Experimentation.....	73
2.3.9	A Lack of Frameworks for Identifying, Measuring, or Eliminating Bias in the News	74
2.3.10	Deductions from the State of the Art on Bias	74
2.3.11	Experiments on the Impact of News Website Design on Perceived Bias.....	74

2.3.12	Conclusions on Bias	75
2.3.13	Summary on Bias	75
2.4	Trust	76
2.4.1	Introduction	76
2.4.2	The Impact of Website Design on Trust	77
2.4.3	Theoretical and Empirical Models, Theories and Frameworks for Trust	79
2.5	Conclusions on the State of the Art	79
2.6	Summary on State of the Art	80
3	CAFE - An Experiment Framework and Platform for the Measurement of Perceived Credibility, or any of its Dimensions or Measures	81
3.1	Introduction.....	81
3.2	Motivation.....	81
3.3	The CAFE Framework	81
3.3.1	An Open Framework.....	83
3.3.2	Framework Description, Objectives, Use, and Purpose.....	83
3.3.2.1	Phases	83
3.3.2.2	Stages.....	85
3.3.2.3	Considerations	86
3.3.2.4	Impact	86
3.3.3	Related Frameworks.....	87
3.4	Benefits of Using the CAFE Framework and Experiment Platform	87
3.5	An Accompanying Experiment Platform	88
3.5.1	Supported Experiment Designs.....	88
3.5.2	Description of Default Functionality	88
3.5.3	Development Technologies.....	89
3.5.4	Extensibility	90
3.5.5	Documentation	90
3.6	Best Practices, Recommendations, and Novel Contributions to the Typical Credibility Experiment Lifecycle	90
3.6.1	Statistical Power Analysis and Effect Size	91
3.6.2	Crowdsourcing Participants	91
3.6.3	A Repository and Classification of Credibility Measures Used in Peer Reviewed Research	91
3.6.3.1	Description	92
3.6.3.2	The Difficulty of Choosing Appropriate Measures.....	92
3.6.3.3	Inclusion in the Repository.....	96
3.6.3.4	Using the Repository	96

3.6.3.5	An Ongoing Classification Task	96
3.6.4	VAS: An Improved Measurement Scale.....	97
3.6.5	The Experiment Matrix, Latin Cubes, and Cuboids	98
3.6.5.1	Balance in the Experiment Design	100
3.6.5.2	Paths Through a Latin Cube Experiment Matrix	101
3.6.5.3	Steps to Increase Experiment Validity in a Latin Cube or Cuboid	103
3.6.5.4	Accounting for the Impact of Textual or Visual Content on Experiments Focusing on the Impact Website Design or Visual Presentation on Perceived Credibility	104
3.6.6	A New Method to Derive a Credibility Score	105
3.6.7	Statistical Testing.....	105
3.6.8	Increased Reporting of Individual Measures and their Impact on the Credibility Score using Visual Aids.....	105
3.6.9	Target Audience and Use of the CAFE Framework, Experiment Platform, and Recommendations	106
3.7	Design, Development, and Implementations of the CAFE Framework and Experiment Platform	106
3.8	The CAFE Framework and Platform: Evaluating Case Study Implementations with Different Experiment Configurations	106
3.9	Discussion.....	107
3.10	Conclusion	110
4	Experiment One – Perception of Bias: The Impact of User Characteristics, Types of News Website Design, and Individual Technical or Structural Features	111
4.1	Introduction	111
4.2	Influence from the State of the Art.....	112
4.3	Experiment Objectives and Contributions.....	115
4.3.1	Objectives	115
4.3.2	Contribution	115
4.4	Research Ethics Approval	115
4.5	Use of Website Titles to Describe Results and Findings.....	115
4.6	Experiment Design	116
4.6.1	Experiment Hypothesis.....	116
4.6.1.1	Primary Hypothesis	116
4.6.1.2	Secondary Hypotheses	116
4.6.1.2.1	Individual Features.....	116
4.6.1.2.2	User Characteristics	117
4.6.1.2.3	Category of News Website Design	117
4.7	Methodology.....	117

4.7.1	Experiment Content	117
4.7.1.1	Websites and Webpages	117
4.7.1.2	News Articles	118
4.7.1.3	Pairing of News Webpages and News Articles	119
4.7.2	Distortions	119
4.7.2.1	Website Logos and Branding	122
4.7.3	The Experiment Platform	122
4.7.4	Definition of Bias and Participant Instructions	123
4.7.5	The Five Stages of the Experiment	124
4.7.6	Crowdsourcing	125
4.7.7	Steps to Ensure Validity of Results	125
4.7.8	Statistical Power and Effect Size	126
4.7.9	Measuring Perceived Bias with Visual Analogue Scales (VAS).....	127
4.8	Participant Sample Profile Overview	128
4.9	Statistical Analysis and Results - Initial Bias Rating	128
4.10	Statistical Analysis and Results – Comparative Re-Evaluation Bias Ratings	129
4.10.1	Main Effects – Comparative Re-Evaluation Bias Ratings	129
4.10.2	Simple Main Effects for Distortions – Comparative Re-Evaluation Bias Ratings	130
4.10.2.1	D5	131
4.10.2.2	D7	132
4.10.3	Simple Main Effects for Webpage/Article Combinations – Comparative Re-Evaluation Bias Ratings	133
4.10.4	Most Positively and Negatively Biased	133
4.11	The Impact of News Website Category on The Perception of Bias	134
4.11.1	Traditional Print News Agencies	134
4.11.2	News Magazines	134
4.11.3	International Multi-Format News Agencies.....	134
4.11.4	The Impact of the Overall Design and Visual Aesthetics of a News Webpage on the Perception of Bias	135
4.12	The Impact of User Characteristics on the Perception of Bias	135
4.12.1	Physical User Characteristics	135
4.12.1.1	Gender	135
4.12.1.2	Age	135
4.12.2	Education, Income, and Socioeconomic	136
4.12.2.1	Income	136
4.12.3	Political Ideology and Affiliation.....	137
4.12.3.1	Party Affiliation	137

4.12.4	Internet Usage and News Access Habits	137
4.12.5	Summary of User Characteristics Results.....	137
4.13	Hypothesis Results.....	138
4.13.1	Primary Hypothesis Results	138
4.13.2	Secondary Hypotheses Results	139
4.13.2.1	Individual Features	139
4.13.2.2	User Characteristics.....	139
4.13.2.3	Category of News Website.....	139
4.14	Reflective Questions.....	140
4.14.1	Least Biased Medium	140
4.14.2	Most Noticed Forms of Bias	141
4.14.3	Common Features and Characteristics of a News Webpage, Including the News Article, which can Impact or Convey Bias	141
4.14.3.1	News Article Features Impact on Perceived Bias	141
4.14.3.2	News Website Homepage Features Impact on Perceived Bias	141
4.14.3.3	Characteristics of a News Website’s Design and Visual Presentation Impact on Perceived Bias	142
4.14.3.4	The Impact of a News Website’s Content Focus on Perceived Bias	143
4.14.3.4.1	The Impact of Website and Third-Party Functionality on Perceived Bias	144
4.15	Lessons Learned from Experiment One	145
4.16	Summary of Experiment One Findings	146
4.17	Conclusion	147
5	Experiment Two – The Impact of Increasing and Decreasing the Professionalism of News Webpage Aesthetics on the Perception of Bias	148
5.1	Introduction	148
5.2	Influence from the State of the Art and from Experiment One	148
5.3	Experiment Objectives and Contributions.....	150
5.3.1	Objectives	150
5.3.2	Contribution	151
5.4	Research Ethics Approval	151
5.5	Use of Website Titles to Describe Results and Findings.....	151
5.6	Experiment Design	151
5.6.1	Hypotheses	152
5.6.1.1	Primary Hypotheses	152
5.6.1.2	Secondary Hypotheses	153
5.7	Methodology.....	153
5.7.1	Experiment Content	154

5.7.1.1	Websites and Webpages	154
5.7.1.2	Creating Biased News Articles.....	155
5.7.1.3	Pairing of News Webpages and News Articles	157
5.7.1.4	Images in News Articles.....	158
5.7.2	Professional Website Design and Webpage Aesthetics	158
5.7.3	Distortions	159
5.7.4	Measuring the Success of the Distortion Process	163
5.7.4.1	Website Logos and Branding	164
5.7.5	The Experiment Platform.....	165
5.7.6	Definition of Bias and Participant Instructions.....	165
5.7.7	Five Stages of the Experiment	166
5.7.8	Crowdsourcing.....	166
5.7.9	Steps to Ensure Validity of Results.....	166
5.7.10	Statistical Power and Effect Size	166
5.7.11	Bias Rating Scale - VAS	167
5.8	Participant Profile	167
5.9	Statistical Analysis and Results – Re-Evaluated Bias Rating	167
5.9.1	Re-Evaluated Bias Ratings.....	168
5.9.2	Main Effects.....	168
5.9.3	Simple Main Effects for Distortions	169
5.9.3.1	Guardian	169
5.9.3.2	Telegraph.....	170
5.9.3.3	Independent	170
5.9.3.4	Economist.....	170
5.9.3.5	Spectator	170
5.9.3.6	New Statesman	170
5.9.3.7	Al Jazeera	170
5.9.3.8	BBC	171
5.9.3.9	Reuters.....	171
5.10	Results.....	171
5.10.1	D0.....	174
5.10.2	D1.....	174
5.10.3	D2.....	174
5.10.4	D3.....	174
5.10.5	D4.....	174
5.10.6	D5.....	175
5.10.7	D6.....	175

5.11	Findings from Experiment Results	175
5.11.1	Comparison of D0 Ground Truth to D1 Least Professional	175
5.11.2	The Cumulative Impact of Professional Webpage Aesthetics: D1 to D5 and Vice Versa	176
5.11.3	The Impact of the Amount of Bias in the News Articles.....	176
5.11.4	The Impact of Category of News Website.....	177
5.12	Hypotheses Results.....	177
5.12.1	Primary Hypothesis.....	177
5.12.2	Secondary Hypotheses	178
5.13	Design Implications	179
5.14	Implications for News Consumers.....	180
5.15	Discussion.....	181
5.16	Why is this Happening?.....	183
5.17	Reflective Questions.....	183
5.17.1	Recognition of Experiment Content and Perception of Bias	183
5.17.2	Impact of Design and Aesthetics on the Perception of Bias and Professionalism.....	184
5.17.3	The Identification of Features and Characteristics of a News Website’s Design, Reflected in a News Webpage, that Impact Perceived Professionalism and Bias	185
5.17.3.1	Professionalism	185
5.17.3.2	Bias.....	185
5.17.4	Most Credible and Most Biased News Mediums	185
5.17.5	Most Relied on Dimensions of Credibility When Judging News Online	185
5.18	Lessons Learned from Experiment Two.....	186
5.19	Summary of Experiment Two Findings	187
5.20	Conclusion	187
6	Experiment Three – Increasing and Decreasing the Perception of Bias Through News Website Design and Webpage Aesthetics.....	189
6.1	Introduction	189
6.2	Influence from the State of the Art and Experiments One and Two	189
6.3	Experiment Objectives and Contributions.....	190
6.3.1	Objectives	190
6.3.2	Contribution	191
6.4	Research Ethics Approval	191
6.5	Use of Website Titles to Describe Results and Findings.....	191
6.6	Experiment Design	191
6.6.1	Hypothesis.....	192
6.7	Methodology.....	192
6.7.1	Experiment Content	193

6.7.1.1	Websites and Webpages	193
6.7.1.2	Creating Biased News Articles.....	193
6.7.1.3	Pairing of News Webpages and News Articles	193
6.7.2	High Quality Aesthetics and Good Design	194
6.7.3	Definitions of Bias and Participant Instructions	197
6.7.4	Crowdsourcing	197
6.7.5	Validity, Statistical Power Analysis, and the Bias Rating Scale.....	197
6.7.6	Profile.....	197
6.8	Statistical Analysis and Results	197
6.8.1	Main Effects	198
6.8.2	Simple Main Effects for Distortions	199
6.8.2.1	Guardian	199
6.8.2.2	Telegraph.....	199
6.8.2.3	Independent	200
6.8.2.4	Economist	200
6.8.2.5	Spectator	200
6.8.2.6	New Statesman	200
6.8.2.7	Al Jazeera	200
6.8.2.8	BBC	200
6.8.2.9	Reuters	201
6.9	Results.....	201
6.10	Findings from the Experiment Results	203
6.10.1	Comparing the D1 Controls to their Respective D2 High Quality Aesthetics.....	203
6.10.2	Comparing the D1 Controls to their Respective D3 Low Quality Aesthetics	204
6.10.3	Comparing the D2 High Quality Aesthetics to their Respective D3 Low Quality Aesthetics	204
6.11	Hypotheses Results	205
6.12	Design Implications	206
6.13	Discussion	207
6.14	Lessons Learned from Experiment Three.....	208
6.15	Summary of Experiment Three Findings.....	208
6.16	Conclusions.....	209
7	Summary Findings	210
7.1	State of the Art.....	210
7.1.1	Isolated Research Efforts	210
7.1.2	A Need for a Classification of Biases that Impact the Different Stages of News.....	210

7.1.3	Bias in the Dissemination of News on the WWW has yet to Receive the Attention it has in Traditional Mediums	210
7.1.4	Most Studies on Bias in the News do not Focus on the Content or the Message Therein	211
7.1.5	The Lack of Theoretical or Empirical Models or Frameworks to Explain How Judgements of Bias in News are Arrived at and their Effect, or how to Measure it	211
7.2	Credibility Assessment Framework for Experimentation (CAFE)	211
7.2.1	An Adaptable Framework and Opensource Platform to aid in the Design, Development, and Deployment of Experiments to Measure Credibility or its Dimensions and Measures	211
7.2.2	Eight Existing Best Practices, Recommendations, and Novel Contributions for Experimentation	212
7.3	Experiment One Findings	212
7.3.1	The Impact of Individual Webpage Features, User Characteristics, and News Website Category on the Perception of Bias	213
7.3.2	Reflective Questions Reveal the Characteristics of a Website’s Design or Webpage’s Aesthetics that Impact Perceived Bias the Most	213
7.3.3	Initial Demonstration and Testing of the CAFE Framework and Experiment Platform ...	213
7.4	Experiment Two Findings	213
7.4.1	Demonstration of the Impact of Characteristics of Website Design Reflected in a Webpage’s Aesthetics Impact Perceived Bias	214
7.4.2	Increased Understanding of Bias for a Proposed Model to Explain How Judgements of Bias are Made and their Effect	214
7.5	Experiment Three Findings	214
7.5.1	Confirmation that Characteristics of a Website’s Design, Reflected in a Webpage’s Aesthetics, Influence Perceived Bias in the News Article Contained Therein	214
7.5.2	Demonstration that Perceived Bias can be Predictively Influenced	215
8	Conclusions	216
8.1	Research Objectives and Achievements	217
8.1.1	Objective One: Research the State of the Art on Measuring Bias and what Frameworks are Used	217
8.1.2	Objective Two: To Design, Develop, and Test an Innovative Framework and Platform that Supports a Range of Experimental Designs to Detect and Record User’s Perception of Bias	219
8.1.3	Objective Three: To Identify Whether Features and Characteristics of a Website’s Design, Reflected in a news Webpage’s Aesthetic, Impact Perceived Bias	220
8.1.4	Objective Four: And if so, To Evaluate the Adaption of Such to Predictively Impact the Perception of Bias	221
8.2	Contributions to the State of the Art	221
8.2.1	Major Contribution	222

8.2.2	Minor Contribution: The Experiment Framework and Platform, Including a Repository of Credibility Measures	223
8.2.3	Minor Contribution: An In-depth Review of Existing Frameworks and the Underlying Theory in the Domain.....	223
8.3	Future Work.....	224
8.4	Final Remarks	225
9	Appendices	257
9.1	Appendix A – Chapter One: Bias in the Research	257
9.2	Appendix B – Chapter Two: Definitions of Bias	258
9.2.1.1.1	Definitions of Media Bias	258
9.2.1.1.2	Definitions of Political Bias	261
9.2.1.1.3	Definitions of News Bias	261
9.2.1.1.4	Definitions of Political Bias in News.....	266
9.2.1.1.5	Descriptions of Content or Photograph Bias.....	267
9.2.1.1.6	Definitions by Differentials in Coverage	268
9.2.1.1.7	Categorisation of Definitions of Bias by Related Terms	268
9.2.1.1.8	Algorithmic Definitions of Bias.....	269
9.2.1.1.9	Definitions of Bias in Social Science.....	270
9.3	Appendix C – Chapter Two: Biases in the Production of News	272
9.3.1	Global Influences that Bias the Production of News	273
9.3.2	Domain Influences that Bias the Production of News	274
9.3.3	Organisational Influences that Bias the Production of News.....	275
9.3.4	Journalistic or Editorial Biases that Influence the Production of News.....	278
9.4	Appendix D – Chapter Two: Bias in the Dissemination of News.....	283
9.4.1	Bias in the Dissemination of News on the Radio.....	283
9.4.2	Bias in the Dissemination of News in Print Newspapers.....	284
9.4.3	Bias in the Dissemination of News on Television	285
9.5	Appendix E - Chapter Two: Studies on Bias in the Dissemination of News in Print Newspapers and News Magazines.....	287
9.5.1	Coverage Bias	287
9.5.2	Headline Bias	287
9.5.3	Photograph Bias	288
9.5.4	Frontrunner Bias.....	289
9.5.5	Visual Bias	289
9.6	Appendix F - Chapter Two: Studies on Bias in the Dissemination of News on the Radio	290
9.6.1	Reasons for the Lack of Research into Bias in the Production or Dissemination of News on the Radio.....	290

9.6.2	Partisan and Class Bias in Australian Radio.....	291
9.6.3	Partisan or Pro Government Bias in South African Broadcasting Corporation.....	291
9.6.4	Identifying Bias in the Speech of Radio Broadcasters.....	291
9.6.5	Favourability and Coverage Bias in Radio	291
9.6.6	Lack of Research on Advertising Bias on Radio	292
9.6.7	Incumbency Bias on Radio	292
9.7	Appendix G – Chapter Two: Studies on Bias in the Dissemination of News on Television	294
9.7.1	Presentation Bias in Television News.....	294
9.7.2	Facial Expression Bias in Television News.....	294
9.7.3	Soundbite and Imagebite Bias in Television News	294
9.7.4	Visual Bias in Television News.....	295
9.8	Appendix H – Chapter Two: Trust.....	297
9.8.1	Introduction.....	297
9.8.2	The Impact of Website Design on Trust.....	297
9.8.3	Theoretical Models and Frameworks of Trust.....	300
9.8.4	Empirical Tools and Frameworks for Measuring Trust.....	301
9.8.5	Conclusions on Trust	302
9.8.6	Summary on Trust.....	303
9.9	Appendix I – Chapter Three: Statistical Power and Effect Size.....	304
9.9.1	Statistical Power.....	304
9.9.2	Defining an Appropriate Statistical Power and Effect Size for an Experiment.....	305
9.10	Appendix J – Chapter Three: Crowdsourcing	306
9.11	Appendix K – Chapter Three: Weighted Credibility Score.....	308
9.11.1.1	An Alternative: The Weighted Credibility Score.....	308
9.12	Appendix L – Chapter Three: Choosing and Reporting Statistical Tests.....	311
9.13	Appendix M – Chapter Three: Using Visual Aids to Report Individual Measures Making up a Credibility Score	312
9.14	Appendix N - Chapter Four: Experiment One – Statistical Analysis and Results – Initial Bias Rating	314
9.14.1	Main Effects – Initial Bias Rating	314
9.14.2	Simple Main Effects for Webpage/Article Combinations – Initial Bias Rating.....	315
9.14.2.1	D0.....	315
9.14.2.2	D1.....	315
9.14.2.3	D2.....	315
9.14.2.4	D3.....	315
9.14.2.5	D4.....	316
9.14.2.6	D5.....	316

9.14.2.7	D6	316
9.14.2.8	D7	316
9.14.2.9	D8	316
9.14.3	Simple Main Effects for Distortions – Initial Bias Rating.....	317
9.14.3.1	D5	318
9.14.3.2	D7	318
9.15	Appendix O - Chapter Four: Experiment One – Statistical Analysis and Results – Comparative Re-Evaluated Bias Ratings – Simple Main Effects for Webpage/Article Combinations	319
9.15.1.1	D0	319
9.15.1.2	D1	319
9.15.1.3	D2	320
9.15.1.4	D3	320
9.15.1.5	D4	320
9.15.1.6	D5	320
9.15.1.7	D6	320
9.15.1.8	D7	320
9.15.1.9	D8	321
9.16	Appendix P - Chapter Four: Experiment One – Frequency of Re-Evaluation	322
9.17	Appendix Q – Chapter Four: Impact of Socioeconomic Status on Perceived Bias.....	324
9.17.1.1	Education	324
9.17.1.2	Occupation.....	324
9.17.1.3	Living Arrangements.....	324
9.18	Appendix R – Chapter Four: Experiment One – Elections, Political Leaning and Ideological Viewpoints	326
9.18.1.1	Political Ideology.....	327
9.19	Appendix S – Chapter Four: Experiment One – WWW and News Consumption Habits.....	328
9.19.1.1	WWW Usage.....	328
9.19.1.2	News Mediums.....	328
9.19.1.2.1	Television News.....	329
9.19.1.2.2	Newspapers	329
9.20	Appendix T – Chapter Four: Experiment One – Reflective Questions	331
9.20.1	Least Biased Medium.....	331
9.20.2	Awareness of Certain Types of Bias.....	332
9.20.3	Political Topics that Regularly Receive Biased Coverage.....	333
9.20.4	Features and Characteristics of a News Website and News Article that Convey Bias.....	334
9.20.5	Impact of News Website Content or Focus on the Perception of Bias	337

9.20.6	Impact of Website Functionality, Third Party Functionality, Personalisation, and Advertising on the Perception of Bias	338
9.21	Appendix U - Chapter Five: Experiment Two - Steps to Increase Validity	339
9.22	Appendix V – Chapter Five: Experiment Two – Reflective Questions	340
9.22.1	The Impact of Website Design and Webpage Aesthetics on the Perception of Bias and Professionalism	340
9.22.2	The Features and/or Characteristics of a News Webpage that Impact Perceived Professionalism and Perceived Bias.....	341
9.22.3	Most Credible and Least Biased News Mediums	343
9.22.4	Dimensions of Credibility.....	345
9.23	Appendix W – Chapter Six: Experiment Three – D0, D1, D2 and D3 Images.....	346
9.24	Appendix X – Chapter Five and Six: Experiment Two and Three - Participant Sample Profile	
	350	

Table of Figures

Figure 1-1 Common dimensions and measures of credibility, which dimensions depends on whether it is the source, medium or message that is being measured and the situation it is being measured in.18

Figure 2-1 A breakdown by Choi and Stvilia of the underlying dimensions of credibility as defined in interpersonal communication and the name of the researchers who defined them (W. Choi & Stvilia, 2015b).....25

Figure 2-2 A depiction of the process of measuring credibility and how a credibility score is derived. ...26

Figure 2-3 The Elaboration Likelihood Model of Persuasion (ELM), adapted from (Petty, 1977; Petty & Cacioppo, 1978, 1981, 1986).31

Figure 2-4 Visual representation of Chaiken’s Heuristic Systematic Model of Information Processing (HSM) which is described, but not visualised, in the paper “Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion” (Chaiken, 1980).33

Figure 3-1 The Credibility Assessment Framework for Experimentation (CAFE).....84

Figure 3-2 Screenshot of part of the repository showing the individual measures on the left, and the domains they are used in at the top93

Figure 3-3 Closeup of the repository shown in Figure 3-2 showing individual measures on the left and the domains they were used in at the top94

Figure 3-4 Symbol key for the codes in the repository.....94

Figure 3-5 Key for the studies shown in the repository with a link to the PDF of each.....95

Figure 3-6 Mono-directional Visual Analogue Scale (VAS) to detect the amount of Bias, 0 – 100, in a news article.97

Figure 3-7 Isometric projection showing path 1 of 36 paths a participant might be assigned to in a 6x6x6 Latin Cube experiment. Each box represents a Website x Distortion x Measure (WDM) combination.99

Figure 3-8 Alternative matrix representation of the 6x6x6 Latin cube shown in Figure 3-7.103

Figure 4-1 A webpage from Independent.co.uk with the main categories of technical features highlighted in colour..... 114

Figure 4-2 Partial screen captures from the Economist website. The top is the D0 (the control), the bottom is the D1 distortion in which all explicit advertising has been removed..... 121

Figure 4-3 Bi-directional Visual Analogue Scale (VAS) from -100 to +100..... 128

Figure 4-4 The most negative on the left – the most positive on the right. 133

Figure 5-1 The Designed Bias Rating, Pre-Test Bias Rating, and D0 Ground Truth Bias Rating from Table 5-1..... 157

Figure 5-2 The D6, D5, D4, D3, D2, D1, and D0 distorted webpage/article combinations for Al Jazeera. 162

Figure 5-3 The D6, D5, D4, D3, D2, and D1 distorted webpage/article combinations for Al Jazeera. ... 163

Figure 5-4 Mono-directional Visual Analogue Scale (VAS) from 0 - 100 used in this experiment. 167

Figure 5-5 A graph showing the Mean bias scores for the nine distorted webpage/article combinations in Table 5-5.	173
Figure 5-6 The Telegraph, Spectator and Independent webpage/article combinations from Figure 5-5. The square boxes indicate significant differences between distortions as per Table 5-5.....	173
Figure 6-1 Breakdown of experiments two and three showing how they utilised the same webpage/article combinations but that for the most part different distortions were compared in each experiment. ..	193
Figure 6-2 Four distortions of the BBC webpage/article combination. Note: see Figure 9-10, Figure 9-11, Figure 9-12, and Figure 9-13 in section 9.23 of the appendix for larger versions.	196
Figure 6-3 The impact of increasing and decreasing the visual quality of the aesthetics for each webpage/article combination on perceived bias.	203
Figure 9-1 A conceptual framework by Hammersley and Gomm which categorises different types of bias in social science research as different types of error (Hammersley and Gomm, 1997).....	271
Figure 9-2 Causes and Forms of Media Bias, from (Park et al. 2009).	272
Figure 9-3 Examples of chyrons from CNN and MSNBC.	286
Figure 9-4 Example of an XY Scatter Graph showing the results from an example credibility evaluation experiment with six common measures of credibility which participants (n = 300) rated using VAS scales from 0-100.	313
Figure 9-5 Political ideology of participants.	326
Figure 9-6 Election participations of participants.....	326
Figure 9-7 Most common mediums participants use to access news.	328
Figure 9-8 US Television news access habits of participants.....	329
Figure 9-9 Newspaper readership of participants.	329
Figure 9-10 The top half of the D0 and D1 webpage/article combinations used in experiment three.	346
Figure 9-11 The bottom half of the D0 and D1 webpage/article combinations used in experiment three.	347
Figure 9-12 The top half of the D2 and D3 webpage/article combinations used in experiment three.	348
Figure 9-13 The bottom half of the D2 and D3 webpage/article combinations used in experiment three.	349

Table of Tables

Table 1-1 The Chapters of this thesis, Research Objectives, Research Contributions, and Research Methodologies.	16
Table 2-1 Ten human-centric Models, Theories, Frameworks, and Schematics (MTFS) explaining how users form judgements of credibility on the WWW.....	40
Table 2-2 Examples of Domain, Agent, Event, and Supertype categorisations of bias.	64
Table 2-3 Categorisation of biases that have been investigated in the dissemination of news on the WWW, their manifestation and effect, including references to relevant papers where found.	67
Table 4-1 Results of re-evaluated One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions test using the Bonferroni correction for multiple comparisons.	132
Table 4-2 Impact of participant’s age on the perception of bias.....	136
Table 4-3 Income levels and the perception of bias.....	136
Table 4-4 Political party affinity and perception of bias.	137
Table 4-5 Regularity of biased reporting in the news on common divisive issues.....	141
Table 5-1 The Article Titles / Headline, Designed Bias Rating, the Pre-Test Bias Rating, the Website each Article was Assigned to, and the D0 Control Ground Truth Bias Rating for each webpage/article combination from the crowdsourced experiment.....	156
Table 5-2 Depicts the seven main factors of professional news website design resulting from the review, the main underlying design or technical means of conveying it, and the action required in the distortion.....	160
Table 5-3 The stages of the cumulative distortion process to reduce the professionalism of the webpage/article combinations.	161
Table 5-4 Cronbach’s Alpha Scores for each webpage/article combination. Pre-test participants were tasked with arranging them from D5 to D1.....	164
Table 5-5 Results of One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions with the Bonferroni correction for multiple comparisons.	172
Table 5-6 Results of two reflective questions asking participants which news websites they recognised in the study and which they thought were biased.....	184
Table 6-1 The results of the news website design review and the resulting distortion process for each of the distorted webpage/article combinations.	195
Table 6-2 Results of nine, One-Way Repeated Measures ANOVA's using the Bonferroni correction for multiple comparisons.	202
Table 9-1 Use of relied upon terms to define bias. The centre column denotes whether indications of time, and the term, that have been used in the authors definition.	269
Table 9-2 Global influences which may bias the production of news.	274
Table 9-3 Domain influences which may bias the production of news.	275
Table 9-4 Organisational influences which may bias the production of news.	278

Table 9-5 Journalistic or editorial influences which may bias the production of news.....	281
Table 9-6 Categorisations of bias that have been investigated in the dissemination of news on radio, their manifestation and effect, including references to relevant papers where found.	283
Table 9-7 Categorisation of bias investigated in the dissemination of news in print news newspapers and magazines, their manifestation and effect, and references to publications where found.	284
Table 9-8 Categorisations of bias that have been investigated in the dissemination of news on the television, their manifestation and effect, including references to relevant papers where found.	285
Table 9-9 Results of nine One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions using the Bonferroni adjustment for multiple comparisons.	317
Table 9-10 Frequency of re-evaluation of webpage/article combinations from the nine websites.	322
Table 9-11 Frequency of re-evaluation of distortions.	323
Table 9-12 Education level and perception of bias.....	324
Table 9-13 Occupation and perception of bias.	324
Table 9-14 Political ideology and the perception of bias.	327
Table 9-15 News medium access habits and the perception of bias.....	328
Table 9-16 Print newspapers and the perception of bias.	330
Table 9-17 Least biased news medium.....	331
Table 9-18 Awareness of gender bias in online news.	332
Table 9-19 Awareness of political bias in online news.	332
Table 9-20 Awareness of racial bias in online news.	332
Table 9-21 Awareness of climate change bias in online news.	332
Table 9-22 Awareness of health care bias in online news.....	332
Table 9-23 Awareness of religious bias in online news.	333
Table 9-24 Awareness of political party favouritism bias in online news.....	333
Table 9-25 Political topics that regularly receive biased coverage online.	333
Table 9-26 The features of a news article which convey bias.	334
Table 9-27 The features of a news websites homepage which can convey bias.	335
Table 9-28 The characteristics of a news website’s design or a news webpage’s aesthetic that impact perceived bias.....	336
Table 9-29 The impact of news website content or focus on the perception of bias.....	337
Table 9-30 The impact of website functionality on the perception of bias.	338
Table 9-31 The impact of third-party functionality on the perception of bias	338
Table 9-32 The impact of personalisation on the perception of bias.....	338
Table 9-33 The impact of advertising on the perception of bias.	338
Table 9-34 The impact of design and aesthetics on perception of bias in the news article.....	340
Table 9-35 The impact of design and aesthetics on perception of bias in the news organisation.	340
Table 9-36 The impact of supporting content on perception of bias in the news organisation.	340

Table 9-37 The impact of design and aesthetics on perceived professionalism of news organisation.....	340
Table 9-38 The features and/or characteristics of a website’s design or a webpage’s aesthetics, or the article therein, that impact the perception of professionalism.....	341
Table 9-39 The features and/or characteristics of a website’s design or a webpage’s aesthetics, or the article therein, that impact the perception of bias.	342
Table 9-40 Most credible new medium.	343
Table 9-41 Least biased news medium.	344
Table 9-42 Dimensions of credibility most relied on when judging news on the WWW.	345
Table 9-43 Sample profile breakdown of the participants in experiments two and three.	351

Abbreviations

ABC	American Broadcasting Company
AHIP	Automatic Human Information Process
ANOVA	Analysis of Variance
ANSI	Aggregate Slant News Index
APA	American Psychological Association
B2C	Business to Consumer
BBC	British Broadcasting Corporation
BI	Bias Index
CAFE	Credibility Assessment Framework for Experimentation
C&AHIP	Controlled and Automatic Human Information Processes
CHIP	Controlled Human Information Process
CBS	Columbia Broadcasting System
CI	Confidence Intervals
CNN	Cable News Network
CMS	Content Management System
CMYK	Cyan Magenta Yellow Key
CSS	Cascading Style Sheets
DPI	Dots Per Inch
ELM	Elaboration Likelihood Model
FAIR	Findable Accessible Interoperable Reusable
FCC	Federal Communications Commission
GG	Greenhouse-Geisser
HCI	Human Computer Interaction
HF	Huynh-Feldt
HSM	Heuristic Systematic Information Processing Model
HTML	Hypertext Markup Language
ITAR-TASS	Information Telegraph Agency of Russia, Telegraph Agency of Sovereign States
LAMP Stack	Linux Apache MySQL PHP Server setup
LA Times	Los Angeles Times
MTS	Mauchly's Test of Sphericity
MTFS	Models, Theories, Frameworks, and Schematics
MSNBC	Microsoft News Broadcasting Company (Formally NBC)
NBC	National Broadcasting Company
NPR	National Public Radio
NYT	New York Times
PBS	Public Broadcasting Service

PIT	Prominence-Interpretation Theory
RGB	Red Green Blue
RSS	Rich Site Summary / Really Simple Syndication
SABC	South African Broadcasting Corporation
SND	Society for News Design
SPSS	Statistical Package for the Social Sciences
TAM	Technology Acceptance Model
UK	United Kingdom
UPI	United Press International
US	United States
USD	United States Dollars
VAS	Visual Analogue Scale
WEIRD	Western, Educated, Industrialized, Rich and Democratic
W3C	World Wide Web Consortium
WSGI	Web Server Gateway Interface
WWW	World Wide Web

Definitions of Key Terms

The following are definitions for some of the commonly used terms in this thesis.

Website's Design

A news website's visual design is derived from an organisation's branding, which includes its logos, colour schemes and style guidelines. While a website's design also includes information architecture, website structure, navigation, and usability, among many other elements, the research described in this thesis focuses on the impact of the visual presentation of a news website on the news articles contained therein. Therefore, a website's design is conceptualised in this thesis as the overarching visual design applied to a website which includes its layout, colour scheme, and alignment, used to convey an impression of the organisation behind it.

Webpage's Aesthetic

A news webpage's aesthetic is primarily descended from the overall design of a news website. However, other factors such as the content being displayed, integrated third party services such as advertising, personalisation, and even the setup of the monitor and computer the webpage is being viewed on influence the aesthetic. In this thesis, a news webpage's aesthetic is defined as the overall combination of elements including the design, layout, colour schemes, content, and structure, in a single news webpage.

Characteristics of a Website's Design

Characteristics of a website's design are the overarching themes and considerations behind a website's look and feel. Part of the overall design philosophy of a website, they reflect the attitude and values of the organisation behind the website. Individual examples of themes include professional, modern, minimalist, quality, traditional, conservative, busy, gaudy, loud, bold, and active. These may be transmitted through coordinated combinations of branding, colour, layout, the number of elements on a page, the prominence of advertising, professionalism and content of images and even the font. The most obvious and common manifestation of differing design characteristics in news websites is the dichotomy between tabloid and traditional news websites.

Features of a Website's Design

Website features are components of a website or webpages that may include or combine text, images and interactive elements, which are used to convey information or provide a service. They are distinct from message features such as argument quality, readability, expert quotations, and rigor (O'Keefe, 2015). Through their inclusion, omission or distortion, they may increase or decrease the perceived level of

credibility among different users. Examples include search facilities, trending news article navigation, author profile and social media integration, and advertising.

Dimensions of Credibility

As described in section 2.2.6, Dimension of Credibility is a term with three meanings in the literature. First, it is used to describe the elements that make up the construct that is credibility. These may include those derived in interpersonal communication studies such as trustworthiness, expertise, dynamism, goodwill, objectivity, and personal attraction. Second, dimensions of credibility may refer to specific context dependent elements in a certain situation. These include; ‘expertise of the communicator on the subject’, ‘trustworthiness in the website’, or ‘expectation to tell the truth’. Third, the term dimension of credibility is used to describe the ‘dimension of credibility’ being measured, typically; source, medium, or message. However, this could also be defined in the study such as ‘credibility of a website’s sponsor’ to describe the owner or organisation behind the website. While this can be somewhat confusing, it could not be helped as this term is used to describe all three in the literature.

Measures of Credibility

These are the individual measures used in an individual scale, such as Likert Scale or Semantic Differential Scale, to measure credibility in a peer reviewed academic study. In most instances, they exactly or almost exactly reflect the dimensions of credibility as encapsulated by the researcher which they are attempting to measure, e.g. the dimension ‘trustworthiness’ might be measured with a single 7-point scale labelled with the term and anchored with the terms “not at all” to “extremely” (Metzger, Flanagin, & Zwarun, 2003). In more complex studies, they may use multiple measures and scales, which individually might bare less relation to credibility, but the combination of which they argue provide a more accurate overall measurement. Examples include Lowry measuring Trustworthiness, Expertise, and Dynamism as dimensions of credibility with 17, 13, and 6 measures respectively, and Hong using 16, 7-point Likert scales as measure for five common dimensions of credibility, expertise, goodwill, trustworthiness, depth, and fairness (Hong, 2006; P. B. Lowry et al., 2014).

Bias – A Core Dimension and Measure of Credibility

Hovland and his Communication and Attitude Change group in Yale defined trustworthiness and expertise as the first two dimensions of credibility (Hovland et al., 1953, p. 21). Since then, additional dimensions and measures such as goodwill, dynamism, bias, accuracy, fairness, and depth of coverage, among others, have been defined. Bias has been used in a wide range of studies measuring credibility (Abdulla et al., 2002; Chung et al., 2010; Flanagin & Metzger, 2000, 2003, 2007; K. S. Freeman & Spyridakis, 2004, 2009; Greer, 2003; Hong, 2006; Johnson & Kaye, 2000, 2002; Metzger, Flanagin, & Zwarun, 2003; Rains & Karmikel, 2009; Sundar, 1999; Sundar & Nass, 2001). This research focuses on bias due to its importance

as a dimension of credibility when judging news on the WWW. In a large study with 2,684 participants, Fogg et al. found that 11.6% of participants cited information bias as a particular concern when evaluating the credibility of information online. However, this rose to 30.2% when evaluating news online (Fogg et al., 2003).

“The intelligence of that creature known as a crowd is the square root of the number of people in it.”

Terry Pratchett, Jingo

1 Introduction

The purpose of news is to inform, report on and act as a record of events past, as well as to influence the events of the future. The advent of the WWW and smartphones has added a new dimension to the dissemination of news. Whilst news is now more accessible than ever, with notifications of breaking news pushed directly to smartphones, traditional standards of reporting are difficult to maintain due to falling staff numbers and the pressures of publication speed (Arant & Anderson, 2001). News has evolved from the formal dissemination of highly composed content through regulated channels at specific times, to a constant stream of updates on rapidly evolving events. Production, curation and dissemination of serious news content in reputable organisations follows standard procedures. Considerations such as ethics, vetting, gatekeeping, privacy, legal risks and recourse have long been part of journalistic standards and are considered at each step (Frost, 2015; Gladney, 1996). Journalists, news broadcasters and editors from reputable organisations hold a responsibility to their readers, listeners and viewers. They are bound by a code of ethics (Keeble, 2008), industry and statutory organisations that may fine or sanction them (Frost, 2015), and ultimately, are subject to challenge in courts of law (T. Carter et al., 2011). Undoubtedly, even with these checks and balances, biased news has and continues to be produced and published. But the fact remains that in traditional news mediums, each news article, radio bulletin or broadcast is a tangible artefact, attributable and retrievable. In comparison, the quality of news available to consumers online is much more varied. Consequently, consumers need to rely much more on their own judgement. Metzger et al. note: *“the filters and control mechanisms, which formerly served to validate and endorse a rather limited number of information outlets, may not be as effective in this new media environment. Absent such controls, information assessment and verification - core components of source, message, and medium credibility - now often become the responsibility of the media consumer.”* (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003).

Bias, real or fake, conscious or unconscious, perceived or unseen, will always be a factor in news. Reputable journalists, editors and news agencies endeavour to investigate and report news in an unbiased manner. However despite this, individual partisan beliefs, cognitive biases, and errors of judgement can impact their decisions (Donsbach, 2004; Patterson & Donsbagh, 1996). News consumers also hope or believe that the news they consume is a true and accurate telling of the facts or balanced analysis without slant or favour. When news consumers recognise or suspect bias, it can have a detrimental effect on their opinion of, and on their consumption of news from the news agency or source, causing them to seek out alternative unbiased, and thus more credible sources of news in future (Baron, 2006). It should be noted however that many news consumers seek affirmation over information. They prioritise news which aligns with their world views and values over accurate or discrepant news.

Bias is a complex construct without a universally accepted formal definition. This problem is well documented and described by Park et al. who maintain: *“it is intrinsically difficult to objectively define what bias is and so is to measure or correct it”* (Park et al., 2009). Entman, paraphrasing Niven states: *“With all the heat and attention it incites among activists and ordinary citizens, bias is yet to be defined clearly, let alone received much serious empirical attention”* (Entman, 2007; Niven, 2002). On first impression, bias in the news is usually thought of in terms of the textual content in news articles or the spoken words of radio or television broadcasts. Upon further reflection, which news is reported, the amount of coverage, when and how often it is reported, and how it is reported, are just some of the many other ways news is subject to the influence of bias. An extensive compendium of literature from the research community supports this. Research into bias in text includes studies into linguistic bias when reporting race related news stories (Gorham, 2006), using machine learning to detect bias in text (Iyyer et al., 2014), and automated methods to identify biased vocabularies based on political speeches (Krestel et al., 2012). Research into bias in relation to which news is reported and the depth of coverage includes studies into partisan bias in economic news (Larcinese et al., 2011), and studies of the selective coverage of political protest events (McCarthy et al., 1996; Oliver & Maney, 2000). Research into when and how often news is reported has focused on partisan bias in television news networks (Groeling, 2008, 2013a) and the use of machine learning and crowdsourcing to quantify selection and framing bias among major news organisations (Budak et al., 2016). Research into bias on how news is reported includes a longitudinal study of anti-incumbent bias in presidential campaigns (D. T. Lowry & Shidler, 1998), and a number of studies on the facial expressions of television news anchors when they are reporting the news (H. S. Friedman et al., 1980; Miller et al., 2007; Zimmerman, 2013). This small sample serves to demonstrate the range and type of biases studied in news media. Individual cognitive biases affecting the consumption of news within the consumer are out of scope for this research and are therefore not listed or covered in detail in this thesis.

One of the reasons for this broad range is the subjective nature of the term ‘bias’. On its own, it can be hard to define. As Doll and Bradley state: *“the term is abstract; therefore, it tends to mean, like ‘democracy,’ what people want it to mean. At worst, bias becomes ‘what I say it is.’”* (Doll & Bradley, 1974). Further evidence of this can be seen in the difficulty of describing or defining bias without putting it in context. At its core, bias in the news can be described as deliberate or accidental imbalance or slant, usually over a period of time (Entman, 2007). News, due to its very nature, can be rife with deliberate or accidental imbalance and slant. It can be implicit and explicit, physical and psychological. In the context of news, it is most commonly found in the deliberate favouring of, or opposition to an individual, organisation, position or argument. This can impact any of the three main stages of the news pipeline, production, dissemination, and/or consumption.

In traditional news mediums, production biases, whether journalistic, editorial or agency, include but are not limited to presentation; source; partisan; selection; agenda setting; coverage; framing; commercial; and bad news bias. Each of these have been widely covered in journalism literature (Barrett & Barrington, 2005; S. DellaVigna & Kaplan, 2006; Entman, 2007; Groeling, 2008; Latham, 2013). The physical manifestations of such biases include but are not limited to headlines; word choice or linguistic bias; article slant; selective quoting and editing; selective use or misuse of facts; imagery; and even the use of appropriate names and personal titles. Dissemination biases include: the amount of coverage; attention; length; time; detail; repetition; position; and the prominence of an article or news report. It can manifest in the framing, layout/design, graphics, photographs, depth of coverage, tone, the level of respect given to an individual or subject, and the grouping of articles and headlines to form a spread in a newspaper or magazine. Consumption or cognitive biases in the audience have also received significant treatment within cognitive psychology literature (Chiang & Knight, 2011; Ehrlinger et al., 2005; Hernandez & Preston, 2013). These include anchoring; loss aversion; confirmation; salience; selective perception; conservatism; endorsements; recency; stereotyping; choice support; blind spot; bandwagon; narrative; and outcome biases. It should be noted that there are significant overlaps between the manifestation and effect of production, dissemination, and consumption biases.

The majority of this research has been undertaken within traditional mediums, print, radio and television, with a significant proportion focusing on the visual presentation, spoken word, audio-visual word, or other means of dissemination. Yet there are no examples of studies investigating the impact of the visual presentation of news websites on perceived bias in the news articles they contain. The research described in this thesis addresses this issue. It begins with an investigation of the impact of features of a website's design on the perception of bias. Features of a website's design are defined as components of websites or webpages that may include or combine text, images and interactive elements, which are used to convey information or provide a service. They include advertising, search facilities, navigation etc. The research described in this thesis also investigated the overarching characteristics of a website's design. Characteristics are defined as the overarching themes and considerations behind a website's look and feel. They include modern, minimalist, quality, traditional, conservative, busy and/or active. More comprehensive definitions and examples of both are provided in the Definitions of Key Terms in the preface and later in their respective experiment chapters 4 (Features), and 5 and 6 (Characteristics).

Credibility is the overarching domain in which this research on bias is conducted within. A detailed review of this overarching domain and a classification of definitions are provided in sections 2.2 and 2.2.4 of the State of the Art. The simplest and one of the most highly cited definitions is that of Fogg and Tseng who maintain: "*Simply put, credibility can be defined as believability. Credible people are believable people; credible information is believable information.*" (Fogg & Tseng, 1999). Consumers choose their news sources based on a number of factors, including its credibility. Fogg and Tseng also state: "*Virtually all*

credibility scholars describe credibility as- [] a perceived quality [] made up of multiple dimensions". Bias, along with trustworthiness, expertise, accuracy, fairness, goodwill, dynamism etc. are core dimensions and measures of credibility. In fact, research has shown that bias is an increasingly important dimension of credibility when judging news, especially online (Fogg et al., 2003). The focus of this thesis is on the impact of the visual presentation of news websites on the perception of bias in the news articles they contain. A wide range of research has previously looked at the impact of the visual presentation of websites, including news websites, on perceived credibility. While most of these studies included bias as a dimension and measure of credibility, none of them have looked at it specifically. Due to the importance of bias as a dimension and measure of credibility, which increases when users are judging news online, it warrants its own investigation, especially as online news continues to take market share from more traditional mediums.

This research contends, and will demonstrate, that the visual presentation of news on the World Wide Web (WWW) can influence the perception of bias in news articles. What's more, the research presented here will demonstrate that perceived bias can be deliberately and predictively influenced based on distortions made to a news webpage's design and aesthetic. It will demonstrate that bias is a core dimension of credibility when users are judging news online. A website's credibility has previously been shown to be impacted not only by structural features of the design (Chung et al., 2012; Flanagin & Metzger, 2007; Rains & Karmikel, 2009; Rieh & Hilligoss, 2008), but also the level of aesthetic treatment (Robins & Holmes, 2008). The studies conducted as part of this thesis have demonstrated that perceived bias is also similarly affected. Robins and Holmes have shown that users judge a website's credibility in an average of 3.42 seconds (Robins & Holmes, 2008). This indicates that users initially rely on heuristic means rather than a systematic, content analysis approach when judging websites. Research in rapid cognition also backs up this rapid judgement of information based on its visual presentation. Lindgaard et al. has shown that users can take as little as 50ms to make major judgements about the acceptability of a website as an information source (Lindgaard et al., 2006). McKnight and Kacmar also demonstrated the importance of first impressions of a website on judgements of credibility, including its perceived quality, reputation, and the user's willingness to explore (McKnight & Kacmar, 2007).

1.1 Motivation

Bias in news has been a constant factor since time immemorial. Several studies have demonstrated its deep-rooted effects (S. DellaVigna & Kaplan, 2006; Gerber et al., 2009; Groseclose, 2011; Morris, 2005). It is considered among the worst transgressions of reputable journalists, editors and news agencies. Bias, or the claim of bias, is also among the first accusations and counter claims by the consumers of news when they encounter opposing opinions or counteracting facts. As one of the pioneers in the field Edith Efron began her book on bias in television network coverage of the 1972 United States (US) presidential election: "*Bias*' is a concept which by now has become a loaded code-word — used as automatic invective by people who dislike the networks on political grounds and denied by those who are politically sympathetic to the

networks, with few of the critics or defenders capable of saying what it is they are talking about” (Efron, 1971, p. 1). Hofstetter also has a similar conclusion in his description: *“If a viewer concludes that a departure from truth is due to incompetence or laziness, he calls the newscaster an ass and labels his output as rubbish. If he concludes that the news-bearer was aware of an untruth, the viewer calls the journalist a liar and says that the news service is shot through with bias.”* (Hofstetter, 1976, p. 4). Stevenson and Green also comment on this issue with: *“What readers consider bias may simply be discrepant information which evokes evaluative response.”* (Stevenson & Greene, 1980). Groeling maintains that this is an issue for researchers investigating bias. When writing about the problem of conducting research in the domain, he stated: *“The problem is exacerbated by the tendency of partisans to strategically allege bias.”* (Groeling, 2013a). Each stage of the news pipeline is rife with opportunities for it’s the conscious or unconscious introduction and influence of bias (Park et al., 2009). The number of opportunities and their effects are increased and exacerbated online among partisan or fringe news agencies, due to lower standards of journalism, less editorial oversight and where issues with legal jurisdiction often prevent or inhibit recourse. It is also an unfortunate fact that those who have biased opinions seek biased information that supports their positions. Thus, many news agencies consciously bias the news they produce, whether, political, environmental or ideological, as a deliberate appeal to the audiences or segment of the market they wish to attract (Gentzkow & Shapiro, 2006; Mullainathan & Shleifer, 2005). However, for reputable mainstream news organisations, being perceived as biased is the antithesis of how they want their target audience to view them. Uncovering whether a news website’s design or a webpage’s aesthetic influences consumer’s perception of bias will enable reputable news agencies to design and/or adapt news websites to reduce the perception of bias.

Credibility has been studied as part of Persuasive Science since Hovland and his team in Yale attempted to develop a unifying theory of credibility (Hovland, 1948; Hovland et al., 1949). Fogg and Tseng define it as: *“Simply put, credibility can be defined as believability”* (Fogg & Tseng, 1999). Credibility is of fundamental importance to journalism and a major factor in the success or failure of a news source (Schweiger, 2000). While McKnight and Kacmar maintain that: *“Website success hinges on how credible the consumers consider the information on the website”* (McKnight & Kacmar, 2007). It emanates from the organisation, board of directors, editors, and individual journalists.

Credibility is a major concern when designing or developing a news agency’s brand or reputation. Schwinger maintains that credibility is an important factor and heuristic in users’ judgements of the acceptability of information sources during times of information overload (Schweiger, 1998). It is considered by news agencies, editors, and journalists when writing news articles, creating broadcasts, radio news segments, or investigative affairs programs. While message features such as argument quality, message quality, repetition, and balance play a major role in credibility evaluations (Cacioppo & Petty, 1989; Chu & Kamal, 2008; Petty & Cacioppo, 1984; Updegraff et al., 2007), dissemination factors such as

visual or aural presentation also play an important role (Burgoon, 1978; Hahn & Cummins, 2014; Weibel et al., 2008). A wide range of factors in the dissemination of news in traditional mediums influence consumers' perception of credibility including: a newspaper's name, brand, logo, layout, the type, style and word choice in headings, and the style and content of images (Fichter & Jonas, 2008). In radio, factors include: the station and show's reputation, and the presenter's interview style. Voice characteristics such as rate, pitch level, pitch verity, volume, fluency, articulation, voice quality, tone, prosody, and cadence, have all been identified as factors which influence credibility (Addington, 1971; Burgoon, 1978; Warhurst et al., 2013). In television, factors include: the network, opening scenes, show style, interview style, and the presenter's style, including their appearance, attractiveness, clothes, manner, posture, eye contact, demeanour, (Brann & Himes, 2010), age (Weibel et al., 2008), vocal characteristics (Bassett et al., 1979; Burgoon & Saine, 1978; O'Neal & Lapitsky, 1991), and even the effect of camera angle (McCain et al., 1977). In each of the traditional mediums, the factors are the means of transmitting the message from the communicator to the consumer during its dissemination. This is also true on the WWW. The means of transmitting the intended message also impact consumers' perception of credibility. This includes the structural features of a webpage and the wider characteristics of a website's design which are visible in a webpage's aesthetic. The level of aesthetic treatment has been shown to impact perceived credibility (Robins & Holmes, 2008). Sundar maintains that it is the underlying cues contained in the design which act as a repository of technical affordances which can be used to judge the credibility of a website (Sundar, 2008; Sundar et al., 2007). Credibility is also a core concern of the consumers of news. It is a factor in their choice of medium, agency and even individual journalists. Fogg et al. has demonstrated that bias is a core dimension and measure of credibility when consumers are judging information online, and that its importance as a measure increases markedly when judging news online (Fogg et al., 2003). Fico, Richardson, & Edwards have shown that as perceived bias in a news article increases, perceived credibility of the news agency behind it decreases (Fico et al., 2004). Consequently, it was decided that the research described in this thesis should focus on whether perceived bias in a news article on the WWW is impacted by its visual presentation and whether this impact can be predictively influenced. This was based on the supposition that perceived bias in news websites is, at least partially, judged heuristically¹, as is done with credibility and some of its other dimensions such as trust.

¹ In this thesis, heuristic judgements mean that users make judgements (at least initially) about bias in the content of a news article, or the message contained therein, based on visual cues in or surrounding the article or in the news webpage. Some of these heuristic judgements may relate to an initial or abstract evaluation of the text or other content of an article such as its length, structure, format, use of inline links or supporting references, the use of data, graphics and images, and even the use of headlines including their choice of words and their gaudiness. This has previously been shown to be the case in the overarching domain of credibility, of which bias is a core dimension and measure.

Reputable² mainstream³ news organisations, editors, and journalists invest a lot of time, effort and money into producing high quality news. Articles are pitched to or assigned by an editor before going through several stages, including research and investigation, fact checking, interviewing, and requesting comments, before final writing and verification (Klinenberg, 2005; Van Hout & Jacobs, 2008). At each stage, factors such as ethics, privacy, gatekeeping and the law are considered, weighed and balanced. An editor may reject a piece or accept it, with an option of commissioning accompanying artwork or photographs. They may also solicit commentary, commission an opposing or devil's advocate article, or an opinion piece to add context to a debate. The position, size and prominence of articles are also considered. Several studies have attempted to define or measure the characteristics of 'good journalism' and/or 'good reporting'. Common characteristics include: objectivity, fairness, balance, unbiased, accuracy, and offering different points of view, while also being thorough and critical (Donsbach & Klett, 1993; Heider et al., 2005). News agencies and individual journalists and editors' reputations are built on years of such practices. The journalistic and editorial policies are set by an editorial board, board of management, trustees and/or owner. Of course, more light natured journalistic pieces and time sensitive contributions may go through an alternative or compressed process, but the fact remains that the production of high-quality news is a long, time consuming, expensive and anti-social process which is often stressful and thankless. With this level of investment, news agencies may be surprised that the news articles they produce could be perceived as more or less biased due to their visual presentation.

Another motivating factor behind the research described in this thesis is the fact that the design and presentation of news websites may deliberately or accidentally be impacting news and therefore influencing public opinion. One of the cornerstones of democracy and a free society is a free press and high-quality news. This has recently been thrown into stark reveal due to the orchestrated campaign by official and unofficial state actors to influence public opinion in recent US elections (Berghel, 2017; Persily, 2017), the #BlackLivesMatter movement (Stewart et al., 2018), and the UK Brexit referendum (Howard & Kollanyi, 2016). It is not surprising that much of these campaigns were fought through social media, a medium that is the fastest growing source of news for younger users.

Increasing polarisation in the news and among audiences in the US creates a widening rift between mainstream liberal and conservative news whilst squeezing the middle ground (Gottfried et al., 2017). While liberals have been shown to trust a wider range of news agencies than conservatives, their distrust,

² We define reputable as: media outlets, editors, and journalists who have a long-standing reputation for reporting factual information without omission. This definition is influenced by Saez-Trumper et al. who maintain that: "*reputable news media are expected to be objective in which stories they report and how they report them; their role is to inform people about what is happening either locally, nationally or worldwide*" (Saez-Trumper et al., 2013), and by Cuthbert who highlights the UNDP/UNESCO quasi-definition: "*a reputable news agency is only concerned with the reporting of fully substantiated, well-sourced or eye-witnessed facts, as well as attributable conclusions or comment, always endeavouring to balance the latter whenever they contain.*" (Cuthbert, 1976).

³ This thesis follows Ellis's definition of mainstream as: "*media that continue to have a generalised distribution of news, comment and analysis to significantly sized audiences and which continue to follow established institutional practices and exert political influence.*" (Ellis, 2014).

once bestowed, is often much stronger (Mitchell et al., 2014). Reacting to, or feeding off this shift in attitudes, many news agencies now have less diversity in opinion among their journalists and editors on topics such as the environment, politics, or science. Two recent studies demonstrate that when opinion is divided on topics, news agencies slant their news to target particular segments of the market (Gentzkow & Shapiro, 2006; Mullainathan & Shleifer, 2005). Increasingly, news agencies present a broad liberal or conservative view with little or reduced dissenting opinions from within the same news agency. This is more obvious in fringe left-wing and right-wing news agencies. This has resulted in increased bias within certain sectors of the US news industry (Groseclose & Milyo, 2005a; Iyengar & Hahn, 2009; Larcinese et al., 2011). Whether this is the cause of or the effect of increased polarisation in the audience in the US is debatable. However, increased bias has been shown to be mirrored within the habits of those viewing or searching for political information online (Allcott & Gentzkow, 2017; Boxell et al., 2017; Garrett & Stroud, 2014). The research described in this thesis was partially motivated to see if the perception of bias in the content could be deliberately influenced through its visual presentation. This is an important motivating factor because it may be useful in demonstrating that certain fringe news organisations make their content appear less biased, and therefore more credible to their target audience, by adapting the design to their individual audiences. Kenix has previously demonstrated that alternative and mainstream news organisations visually represent themselves differently to attract their target audience (Kenix, 2013).

Another motivating factor for this research is the fact that the WWW as a source of news is becoming increasingly popular. Therefore, the impetus to conduct this research and similar work is also growing. In recent years, several major newspapers have decided to cease print publications. Whilst, the predicted immediate demise of traditional news mediums proved unfounded, or at least too imminent, the fact remains that with the continual decline in sales of print, falling radio listenership and network news viewership, that the WWW is becoming increasingly important as a news dissemination medium.

1.2 Research Question and Objectives

1.2.1 Research Question

The question that the research described in this thesis seeks to answer is:

“Can the features and characteristics of a news website’s design, reflected in a news webpage’s aesthetic, influence users’ perception of bias, and if so, to what extent can their deliberate use reasonably predict that perception of bias?”

1.2.2 Research Objectives

To address the research question outlined above, the following research objectives were devised for this thesis:

- **Objective 1:** To research the State of the Art on measuring bias and identify the frameworks and processes which have been used.
- **Objective 2:** To research and design an innovative framework and implement a platform that supports a range of experimental designs to detect and record users' perception of bias.
- **Objective 3:** To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impact perceived bias.
- **Objective 4:** If so, to investigate if the adaptation of such features and characteristics can be used to predictively influence the perception of bias.

Objective 1: To research the State of the Art on measuring bias and identify the frameworks and processes which have been used.

This objective aims to identify existing empirical frameworks for measuring bias, specifically in online news, but also in other mediums. A secondary aim of the review is to identify any theoretical frameworks which may help inform the process. Should the review fail to identify suitable theoretical and/or empirical frameworks, aligned and/or overarching domains such as Trust or Credibility will be investigated. A final aim of the objective is to identify and review any underpinning theory as to how such judgements are made.

Objective 2: To research and design an innovative framework and implement a platform that supports a range of experimental designs to detect and record users' perception of bias.

The second objective aims to design a framework to aid researchers in the design of empirical studies aimed at investigating, measuring, or comparing perceived bias. It would also require the development of a platform to aid researchers in the deployment of such experiments. This would help novice researchers, those new to the domain, or those without the required technical ability, to quickly design, develop, and deploy experiments to measure bias online and in other mediums.

Objective 3: To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impacts perceived bias.

One of the core underpinning suppositions behind the research described in this thesis is that features and characteristics of a news website's design or a news webpage's aesthetic impact perceptions of bias of the news articles contained therein. This is based on the knowledge that website design has previously been shown to impact perceived credibility, and that bias is a core dimension and measure of credibility, especially when judging news online. Thus, it stands to reason that when users are judging news online, the news website's design and presentation could impact their perception of bias.

Objective 4: If so, to investigate if the adaptation of such features and characteristics can be used to predictively influence the perception of bias.

The final research objective focuses on evaluating whether features or characteristics of a website's design, reflected in a news webpage's aesthetic, could be predictively adapted to impact perceived bias. This is important for two reasons. First, it would validate the earlier experiment results. Second, being able to predictively influence user's perception of bias in news articles through their visual presentation has far reaching consequences for the domain.

This finding would also support the supposition that the design or visual presentation of news websites may be impacting in unknown ways on the other dimensions and measures of credibility such as trustworthiness, expertise, accuracy, fairness, and depth of coverage. Consequently, the perception of the message in each news article may be subject to known or unknown influences by the design or visual presentation of the news webpage.

1.3 Research Contributions

This work makes notable contributions to the state of the art in the area of credibility, specifically the measurement of user perception of bias, and its relationship to the visual presentation of news content online. This research provides one major and two minor contributions to the field. They are the direct outcome of addressing the research question and the subsequently defined objectives.

- **Major Contribution:** The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.
- **Minor Contribution:** The experiment framework and platform, including a repository of credibility measures.
- **Minor Contribution:** An in-depth review of existing frameworks and the underlying theory in the domain.

Major Contribution: The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

As stated previously, bias has been shown to be a core dimension of credibility, especially when users are judging news online. Perceived credibility has previously been shown to be impacted by structural elements of a website's design and visual presentation. Perceived bias has also been shown to be impacted by its means of dissemination in traditional news mediums. An increase in perceived bias in a news article results in a decrease in perceived credibility of the news agency behind it.

This major contribution is important for five reasons.

- It has thus far received no attention in the literature, except as one of many dimensions and measures of credibility. Due to its importance as a dimension and measure, especially when judging news online, it warrants individual investigation.
- When the perception of bias in a news article increases, there is a corresponding detrimental effect on perceived credibility of the news organisation behind it. Thus, the findings from this thesis are important to journalists, editors, online editors, and news website designers.
- The predictive nature of the effect. The third experiment detailed in this thesis demonstrates that by increasing and decreasing the visual quality of a news webpage's aesthetic, it was possible to decrease and increase the perception of bias in the news articles they contain. This effect was common across all websites and across articles with different levels of perceived bias.
- Whilst the impact of website design has previously been demonstrated with other dimensions and measures of credibility such as trust (see section 9.8.2), other measures such as Accuracy and Expertise have received less attention due to being seen as the preserve of message rather than medium. Consequently, the impact of website design on these measures warrants further investigation, especially in domains such as e-health where such measures have been shown to be core dimensions of credibility.
- This research demonstrates that the means by which news is presented to the consumer may be having undue influence, whether intentionally or unintentionally, on the intended message in the content of a news article.

Minor Contribution: The experiment framework and platform, including a repository of credibility measures.

Originally one of the aims of this research was to create an empirical framework to measure the impact of a website's design on perceived bias. However, this aim was adjusted after the initial investigation revealed a lack of theoretical underpinnings and comparative research frameworks on the subject. The modified research objective therefore focused on the creation of an empirical framework to measure the impact of a website's design on perceived credibility, and any of its individual dimensions, including bias. There were three reasons behind this change. First, with no existing human-centric models, theories or frameworks explaining how consumers of information online make judgements of bias, it was felt that an empirical framework would be somewhat hermitic, or without theoretical support. Conversely, the overarching domain of credibility has ten, human-centric models, theories, frameworks, and schematics explaining how individuals make judgements of credibility online. It should be noted that each of these put forth different explanations as to how users arrive at judgements of credibility online. The framework proposed in this research is an open empirical framework and experiment platform for conducting experiments. This is also supported by an in-depth canon of credibility research under Persuasive Science going back to the 1930s.

This would provide the underpinning theoretical support that an empirical framework would require. Second, and as stated previously, there are almost no comparable experiments investigating the impact of website design on the perception of bias. The few experiments that do investigate bias in the dissemination of news online tend to focus on specific issues such as headline, photograph, presentation or visual biases, see section 2.3.6. Consequently, it would be difficult to base the work in this thesis on existing research in the domain by pointing out issues and inconsistencies in existing experimental practices which would be remedied with a standardised framework. In contrast, credibility has a host of research investigating the impact of design on perceived credibility which could be referenced. Third, by creating an empirical framework and experiment platform for the design, development, and deployment of experiments to measure credibility, or any of its individual measures such as bias, would have more utility in the domain.

Minor Contribution: An in-depth review of existing frameworks and the underlying theory in the domain.

The initial aim was to discover any underlying theory on forming judgements of bias online or in any other news medium. The secondary aim was to ascertain if there are any empirical frameworks for measuring perceived bias which could be adopted for this research. As none were found, the aligned domain of trust was explored for relevant theoretical and empirical frameworks. Trust is another core dimension and measure of credibility (Hovland et al., 1953, p. 21). It was primarily selected due to the fact that it is a large research area and thus more likely to have relevant frameworks. While ten theories and five frameworks were identified explaining how judgements of trust are made and six empirical frameworks describing how it could be measured, these were deemed unsuitable for bias. Consequently, the overarching domain of credibility was explored. This review identified ten, human-centric Models, Theories, Frameworks, and Schematics (MTFS) which explain how users form judgements online. It then details and compares the underlying theory, namely the Elaboration Likelihood Model of Persuasion (ELM), the Heuristic Systematic Model of Information Processing (HSM), and the Controlled and Automatic Human Information Processes (C&AHIP) (Chaiken, 1980; Petty & Cacioppo, 1986; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). However, no empirical frameworks for the measurement of credibility were identified. This raft of theory in this domain was a major contributor to the decision to design and develop the Credibility Assessment Framework for Experimentation (CAFE) to enable researchers to design and develop experiments to measure credibility online, or any of its dimensions and measures such as bias.

1.3.1 Published Conference Papers

The following conference papers have been published:

1. Spillane, B., Lawless, S., & Wade, V. (2017b). Perception of Bias: The Impact of User Characteristics, Website Design and Technical Features. In Proceedings of the International Conference on Web Intelligence (pp. 227–236). New York, NY, USA: ACM. <https://doi.org/10.1145/3106426.3106474>

The main contribution of this paper is a demonstration of the impact of certain individual technical features of a website's design on the perception of bias. It also found that certain user characteristics also impacted how users perceived bias. Most importantly, it found that certain types of news website design, such as those which were backed by a print publication were considered less biased. The distinct visual characteristics of the websites of traditional print newspapers were considered considerably less biased than those of news magazines and those of international multi-format news agencies whose typical main channel of distribution was via television. This study partly satisfies the main contribution of this thesis.

2. Spillane, B., Lawless, S., & Wade, V. (2018). Increasing and Decreasing Perceived Bias by Distorting the Quality of News Website Design. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers*: (p. 10). Belfast, Northern Ireland.
<http://dx.doi.org/10.14236/ewic/HCI2018.61>

This paper demonstrates that it is possible to predictively decrease or increase the perception of bias by increasing or decreasing the visual quality of a website's design or visual presentation. This forms one of the minor contributions of this thesis.

3. Spillane, B., Lawless, S., & Wade, V. (2020, July 19). The Impact of Increasing and Decreasing the Professionalism of News Webpage Aesthetics on the Perception of Bias in News Articles. *Proceedings of the 22nd International Conference On Human-Computer Interaction*. International Conference On Human-Computer Interaction, Copenhagen, Denmark. https://doi.org/10.1007/978-3-030-49059-1_50

This paper demonstrates the cumulative impact of increasingly professional news website presentation on reducing the perception of bias. It also demonstrates that as the professionalism of a news website's visual presentation decreases, the perception of bias in the article increases. This is a serious issue for news agencies. Fico et al. previously demonstrated that as the perception of bias in a news article increases the credibility of the newspaper publishing it decreases (Fico et al., 2004). This study partly satisfies the main contribution of this thesis.

4. Spillane, B., Lawless, S., & Wade, V. (2017a). Measuring Bias in News Websites, Towards a Model for Personalization. In *Proceedings of the 25th Conference on User Modelling, Adaptation and Personalization* (pp. 387–388). New York, NY, USA: ACM.
<https://doi.org/10.1145/3079628.3079647>

This paper focuses on the benefit of the experiment methodology used in experiment one and used subsequently in experiments two and three, including re-evaluated ratings and the experiment process. It also proposed a model for news personalisation based on perceived bias.

5. Spillane, B., & Wade, V. (2020). A Classification of Biases Relating to the Production, Dissemination and Consumption of News. *CHI 2020 Workshop on Detection and Design for Cognitive Biases in People and Computing Systems*, 8.

http://critical-media.org/cobi/docs/Spillane_Bias%20Classification%20News.pdf

This paper provides an overview into ongoing work to create a classification of biases impacting the production, dissemination, and consumption of news.

1.3.2 *Journal Papers being Prepared for Submission*

The following papers are being prepared, in the following order, for submission to appropriate Journals. Each is based on a significant proportion of the work undertaken as part of this thesis.

1. Spillane, B., Lawless, S., & Wade, V. – The Evolution of the ten human-centric, models, theories, frameworks, and schematics of credibility judgement online, and the ELM, HSM, and the C&AHIP which underpin them.

This will include expanded versions of sections 2.2.8 and 0 from this thesis, including the realisation of the HSM, and likely a realisation of the C&AHIP.

2. Spillane, B., Lawless, S., & Wade, V. – Using Visual Analogue Scales (VAS) to Measure Credibility, or its Dimensions and Measures.

This paper will focus on the benefits of using Visual Analogue Scales (VAS) as a superior scale for measuring credibility online or any of its dimension and measures, such as bias, accuracy, fairness, trust or expertise. Currently most researchers typically employ Likert or Likert type scales. The issues with Likert scales and other scales are discussed in detail in sections 2.2.12 of this thesis along with several other issues including the choice of and number of measures and the sourcing of participants. The benefits of VAS are discussed in detail in section 3.6.4. This paper will also include validity and reliability tests of the scales for credibility.

3. Spillane, B., Lawless, S., & Wade, V. - CAFE: An Experimental Framework and Platform for the Measurement of Perceived Credibility, or any of its Dimensions or Measures.

This paper will highlight the benefits of the empirical framework and accompanying opensource platform used in to conduct the experiments undertaken as part of this thesis, which is designed to aid researchers in the design, development, and deployment of experiments to measure credibility online, or any of its individual dimensions and measures. This is discussed in detail in chapter 3. It is primarily aimed at novice researchers, those new to the domain, or those without the technical ability to deploy such experiments. The opensource platform is also undergoing redevelopment to add additional functionality and to simplify deployment.

1.4 Research Methodology

A *formal research approach* rather than an *informal research approach* was utilised to conduct the three experiments that make up this thesis. A *formal approach* is more systematic and structured. It is typically adopted when a study has a specific aim, a research question that can be answered, and/or there is some similar research in the wider domain that can guide it (Garforth et al., 1998). Thus, the experiment techniques follow protocols and existing methodologies already established in the wider domain. As the overall thesis and each experiment were based on an underlying theory and individual hypotheses, a *deductive research mode* was adopted. This is typically associated with the testing of theory and hypothesis, and quantitative research, rather than an *inductive research mode* which is associated with generating new theory and qualitative research (Creswell, 2013). This thesis also embraces a *positivistic research philosophy*, which maintains that knowledge of the domain can be gained by hypotheses or theory testing through quantitative study designs and statistical analysis. Positivism is often considered the opposite or alternative of an *interpretivist philosophy* which focuses on humanistic qualitative methods, though Lee has shown that both can be employed successfully (A. S. Lee, 1991). There are four main categories of quantitative research study designs; *descriptive, correlational, quasi-experimental, and experimental*. An *experimental design* was chosen as the most appropriate test for each of the experiments and their hypotheses (Wisker, 2007). This involved the application of a treatment, in the form of a distortion, to a news webpage containing a news article to ascertain if perceived bias was impacted. Thus, each experiment has two independent variables (webpage/article x distortions), and one dependent variable (bias). Similar approaches are commonly used in related domains such as user experience testing (Albert & Tullis, 2013), or A/B testing of different website designs (Siroker & Koomen, 2013). Each experiment would generate large volumes of continuous interval type data which would be suited to a number of statistical analysis techniques. Each of the experiments would also generate additional survey data in the form of pre experimental sample profile data (age, gender, internet usage etc.) and post experiment reflective question data (how bias affects them, how bias is communicated etc.). Some qualitative data would also be collected in the form of open comment facilities on some of the reflective questions. The above description is guided by Buckley, Buckley, and Chaing's framework for defining research methodology (Buckley et al., 1976).

1.5 Thesis Structure

Table 1-1 depicts the nine chapters in this thesis, the research objectives, research contributions and the research methodologies. This table is organised by chapter to identify where in this thesis each is dealt with. Of course, there is some overlap between chapters and the research objectives and research contributions. It should also be noted that the Findings and Conclusion chapters due to their nature, contribute to each research objective and research contribution.

Table 1-1 The Chapters of this thesis, Research Objectives, Research Contributions, and Research Methodologies.

Chapter	Research Objective	Research Contribution	Research Methodology
1. Introduction	NA	NA	NA
2. State of the Art	Objective 1 To research the State of the Art on measuring bias and identify the frameworks and processes which have been used.	Minor Contribution An in-depth review of existing frameworks and the underlying theory in the domain	Investigative, Explorative, Review, Analytical
3. The Credibility Assessment Framework for Experimentation (CAFE)	Objective 2 To design, develop and test an innovative framework and platform that supports a range of experimental designs to detect and record user's perception of bias	Minor Contribution The Experiment Framework and Platform, including a repository of measures	Evaluation, Investigative, Analytical
4. Experiment One	Objective 3 To identify whether features and characteristics of a website's design, reflected in a news webpage's aesthetic, impact perceived bias	Major Contribution The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.	Experimental, Hypothesis Testing, Survey (pre, and post experiment)
5. Experiment Two			
6. Experiment Three			
7. Findings	All above	All above	NA
8. Conclusions	All above	All above	NA
9. Appendices	All above	All above	NA

1.5.1 Combined Experiments Two and Three

Due to the complexity and cost of recruiting sufficient crowdsourced participants for the experiments, Experiments Two and Three were executed at the same time. From the participants' point of view, it was one large experiment which required them to rate the perceived bias in a series of news articles. For clarity, these are presented separately.

1.6 The Thesis Focus – Bias, as a Dimension and Measure of Credibility

This thesis and the underlying research are focused on bias as a dimension and measure of credibility, see Figure 1-1. Credibility is the overarching domain of which bias is just one of its many interchangeable dimensions. Which dimensions make up the overarching construct of credibility depends on which

dimension of credibility, source, medium, or message, is being measured, and the domain it is being measured in. An example is the importance of bias as a dimension of the credibility for news, whereas trust would be considered a much more important dimension for e-commerce.

A large proportion of the literature review which follows in chapter 2, is in the overarching domain of credibility. There are two main reasons for this. 1) There is almost no comparable research on the impact of website design on the perception of bias. 2) There are no human-centric models, theories or frameworks explaining how individuals judge bias online. Consequently, much of the literature review and one of the main contributions focus on the overarching domain of credibility, where there is a large body of empirical research and a strong underlying theory. The purpose of this section is to firmly establish the frame of reference for this thesis.

Figure 1-1 depicts bias as a core dimension of a multifaceted construct of credibility along with other dimensions such as expertise, accuracy, fairness, trustworthiness and depth of coverage to name but a few. The dimensions are interchangeable depending on whether it is the credibility of source, message or medium that is being measured, and depending on the domain that it is being measured in. Even for very similar conceptualisations, the dimensions may change. The set of dimensions used to measure the credibility of a person as a source, represented by a news anchor, would be different to the set used to measure the credibility of a source such as a news organisation such as the BBC, Reuters, or The Guardian. In the former, dimensions such as 'Professional' or 'Attractiveness' may be appropriate (Chung et al., 2010; Flanagan & Metzger, 2007), while in the latter, researchers may utilise terms such as 'Community Concern' or 'Desire to make profits' (S. A. Banning & Sweetser, 2007; Kioussis, 2001a). There is also no standard number of dimensions used or an agreed method to create a credibility score. The dimensions are typically chosen in one of two ways. 1) From the existing literature or similar experimentation. 2) Based on the results of a survey as the dimensions accurately encapsulating the concept of credibility of source, medium or message, and within a specific domain. In many cases the dimensions are the titles of groupings of similar terms or phrases, based on the results of a factor analysis of the user surveys. After the dimensions of the concept area have been decided, they are then implemented as measures. Once implemented in an experiment, the dimension of credibility becomes a measure of credibility. Therefore, the dimension 'Trustworthiness of the communicator to tell the truth', may be implemented as the measure 'Trust' or 'Trust / Distrust' in some form of measurement scale.

Bias is a core dimension of credibility when judging news on the WWW. In a study with 2,648 participants, Fogg et al. found that 30.2% relied on the dimension when judging news online (Fogg et al., 2003). An even higher percentage, 53.58%, was found in experiment two of this research, to rely on bias as a core

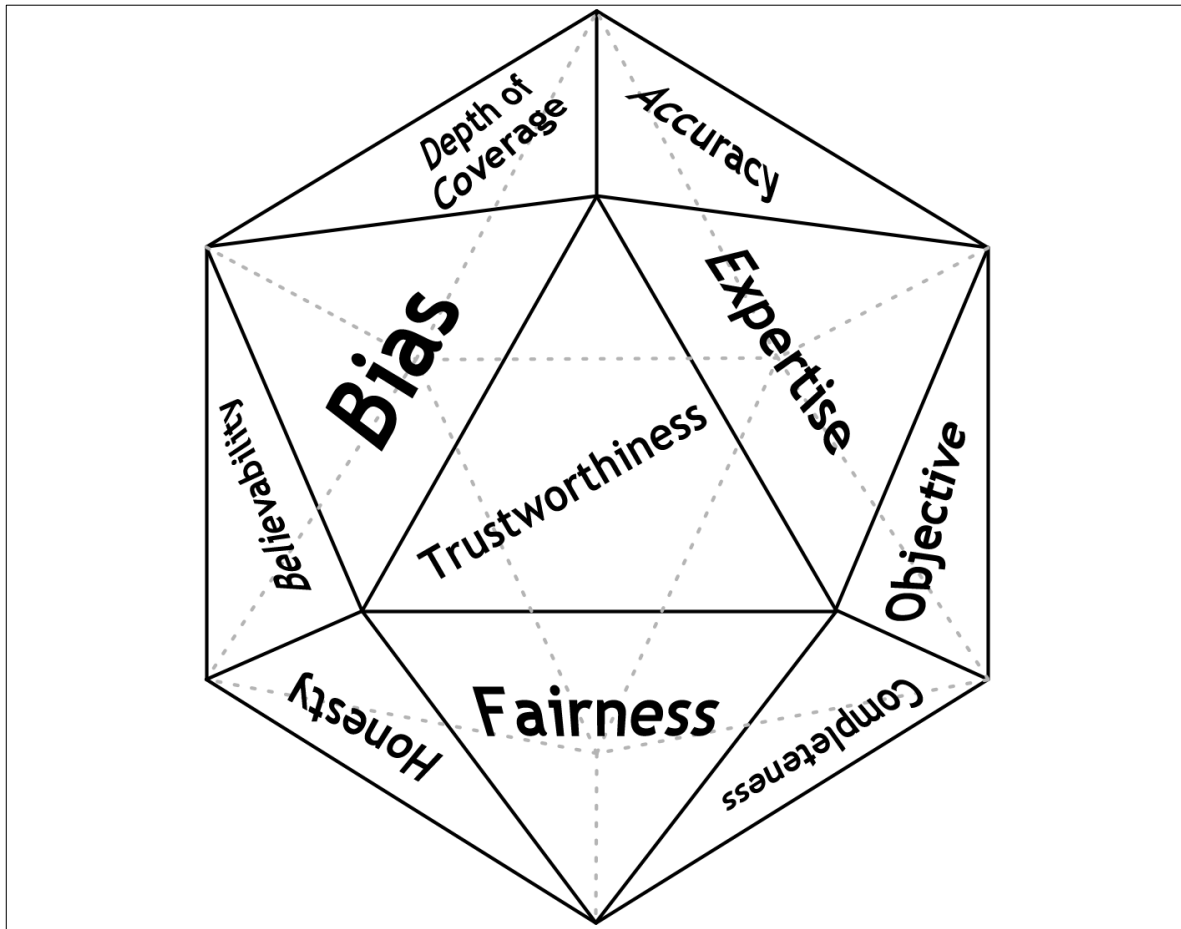


Figure 1-1 Common dimensions and measures of credibility, which dimensions depends on whether it is the source, medium or message that is being measured and the situation it is being measured in.

dimension of credibility (Spillane et al., 2018). Since Hovland and his team in Yale defined the first two dimensions of credibility as trustworthiness and expertise (Hovland et al., 1953, p. 21), a range of additional underlying dimensions such as goodwill, objectivity and dynamism have been defined in interpersonal communication. These are discussed further in section 2.2.6 of this thesis. For empirical investigations into credibility online, researches have adopted some of the dimensions from interpersonal communication, most notably, trustworthiness and expertise, and they have defined specific dimensions such as, bias, fairness, accuracy, depth of coverage etc. Consequently bias has regularly been defined as a core dimension of credibility and used as a measure in a broad range of experiments (Abdulla et al., 2002; Chung et al., 2010; Flanagin & Metzger, 2000, 2003, 2007; K. S. Freeman & Spyridakis, 2004, 2009; Greer, 2003; Hong, 2006; Johnson & Kaye, 2000, 2002; Metzger, Flanagin, & Zwarun, 2003; Rains & Karmikel, 2009; Sundar, 1999; Sundar & Nass, 2001). When this work was initiated, bias quickly became the focus due to its importance to both credibility and news, and because of the large gap in the literature as to whether a news website's design could impact perceived bias.

1.7 Bias in the Research

Despite every precaution taken to prevent its influence one or more biases may exist in this research. There are two overall forms of possible bias, participant bias and researcher bias, these are further discussed in section 9.1 of the appendix.

1.8 Application Domain and Findings

The experiments conducted as part of this thesis focused on news websites and as such all findings are applicable specifically to news websites. This thesis makes no claims that the findings are applicable to other types of websites. While the thesis may refer to “a website” or “websites” upon occasion for brevity or clarity of language, the reader should be cognisant of the specific domain this research was conducted in and in which the claims are made.

"Burke said there were Three Estates in Parliament; but, in the Reporters' Gallery yonder, there sat a Fourth Estate more important far than they all" On Heroes, Hero-Worship, and The Heroic in History,
by Thomas Carlyle, 1841

2 Related Work – State of the Art

2.1 Introduction to State of the Art

The focus of this PhD is detecting whether news website design and webpage aesthetics can impact perceived bias and whether it can be predictively manipulated to that effect. This State of the Art has three main sections. The first focuses on credibility, the overarching research domain. The second is Bias, the main focus of this research. The last is trust, an aligned and comparable research area.

Credibility: The first section of this State of the Art focuses on credibility; the overarching domain of which bias and trust are core dimensions and measures. It begins with definitions of the term before defining the difference between dimensions of credibility and measures of credibility, a longstanding nomenclature issue in the domain. It then demonstrates the importance of bias as a core dimension and measure of credibility, especially when users are judging news online. The State of the Art then highlights ten, human-centric Models, Theories, Frameworks, and Schematics (MTFS), explaining how individuals judge credibility online. This is in direct comparison to the complete lack of theoretical grounding conjecturing as to how bias is judged online. To fully understand the proposed theories, as to how users judge credibility online, the underlying theory provided by the Dual-Process Models of Persuasion is explored. This is underpinned by the Controlled and Automatic Human Information Processes (C&AHIP). It is the contention of this research that users may judge perceived bias in news websites using the peripheral routes or heuristics strategies from these models. The State of the Art then explores existing credibility literature which demonstrates that features and characteristics of a website's design, reflected in a news webpage's aesthetic, are responsible for judgements about the acceptability of a website as an information source. Following this the thesis highlights existing issues with current experimentation practices. Lastly, this chapter makes the case for an empirical framework for measuring credibility online, including each of its dimensions, such as bias, which will be the focus of chapter 3.

Bias: The second section begins with a general introduction to bias, before the paucity of underlying research on the bias is highlighted. To provide context, a section highlighting the longitudinal trends and the obvious gaps in the literature is presented. An in-depth review of definitions of the construct is then provided categorised by domain. It is believed that this categorisation is by far the most comprehensive in the literature. An alternative categorisation of definitions by related terms is then offered before an overview of the influence of bias on each stage of the news production pipeline is explored. The literature review then highlights significant contributions on the impact of bias in the dissemination of news in traditional media before focusing on research on the impact of bias in the dissemination of news online.

The few examples demonstrate that the impact of the visual presentation of news on the WWW has been ignored. The State of the Art then demonstrates that there are two significant gaps within the literature, a lack of human-centric frameworks explaining how individuals make judgements of bias online or in any other medium, and a lack of an empirical framework to measure perceived bias. The State of the Art then explores how bias may be transmitted through the presentation of news online and maps out directions for experimentation.

Trust: The third section briefly reviews an aligned domain and another core dimension and measure of credibility, see section 2.4. The primary purpose of this is to demonstrate the comparative lack of research on bias in comparison to the volume of research on trust. The secondary aim was to identify theoretical and empirical frameworks which may inform the design and development of studies investigating bias. While some of the key points and findings are in section 2.4, for brevity, the majority of the review was moved to section 9.8 of the appendix.

2.2 Credibility

This section introduces and explores the overarching domain of credibility, of which bias is a core dimension and measure, especially when judging news online. The primary aim being to identify models, theories and frameworks which could explain how users form judgements of credibility online. The secondary aim is to inform the development of an empirical framework and platform which can be used to measure credibility online, or any of its dimensions and measures, such as bias.

2.2.1 Introduction to Credibility

Credibility has been studied as part of Persuasive Science since at least the 1940s (Hovland, 1948; Hovland et al., 1949). Through their investigation of the influence of the source in forming opinions in the recipient, Hovland and his team defined the first two dimensions, trustworthiness and expertise (Hovland et al., 1953). Comparing the relative credibility of different mediums was regularly investigated as radio and print gave way to the supremacy of television news (Gantz, 1981; J. Newhagen & Nass, 1989), which in turn, has begun to give way to the WWW (Johnson & Kaye, 1998, 2010). Simultaneously work was being undertaken to define the dimensions of the construct itself while attempting to understand how credibility judgments are made (C. Gaziano & McGrath, 1986; P. Meyer, 1988). In turn these dimensions became measures when they were operationalised in an experiment, the scores of which were combined to create an overall credibility score (Chung et al., 2012; Fogg et al., 2001). More recently, a succession of theories have been proposed to explain how judgments of credibility are formed (Fogg, 2003; Fritch & Cromwell, 2001; Hilligoss & Rieh, 2008; Metzger, 2007; Metzger & Flanagin, 2015; Rieh, 2002; Sundar, 2008; Wathen & Burkell, 2002). The literature review undertaken as part of this thesis identified a broad range of issues which were used to define the requirements which motivated the design and development of a standardised, rigorous empirical framework to assess credibility online. These issues have resulted in a

large disparity in the quality of existing research in the domain that investigates features and characteristics of a website's design, reflected in a news webpage's aesthetic, which impact on perceived credibility. Features are individual, often independent elements of a website or webpage which provide a service or convey information. Examples include search facilities, trending news articles, navigation, author profiles, social media integration, comment facilities, and advertising. Characteristics are the overarching themes of the website's design, such as its level of professionalism, busyness, boldness, loudness, or minimalism. Much of the existing research is also quite fragmented with only partial investigations conducted into individual types of website. Comparative research between different genres of websites, and between websites and more traditional information dissemination mediums such as print or television, is also of vastly differing quality. Consequently, it is difficult to make longitudinal comparisons between similar experiments. This research aims to explore these issues and others further and to identify the requirements of a proposed solution.

2.2.2 *Origins of Research into Credibility*

Credibility research in the modern scientific era can be traced back to Hovland and his Communication and Attitude Change group in Yale (Hovland, 1948; Hovland et al., 1949, 1953; Hovland & Weiss, 1951). However, it was Aristotle in his work on rhetoric who undertook the first recorded scholarly endeavour. He proposed three modes of persuasion; Ethos (the ethical appeal), Pathos (the emotional appeal), and Logos (the logical appeal). By using the Ethos or the ethical appeal, the communicator attempts to convince the receiver of the credibility of their character. Aristotle stated: *"There are three things which inspire confidence in the orator's own character-the three, namely, that induce us to believe a thing apart from any proof of it: good sense, good moral character, and goodwill."* (Translation by W. Rhys Roberts) (Aristotle, 2015). By using the ethos appeal, a communicator relies on their characteristics such as their reputation or expertise to convince an audience: *"to believe a thing apart from any proof of it"*. The English word ethic is derived from the Greek word *ēthos*. By defining good sense, good will, and good moral character as inspiring credibility Aristotle also inadvertently defined the first three dimensions of the term.

2.2.3 *Credibility Research on the Internet*

Historically a large proportion of credibility research has focused on comparing different news mediums. With the advent of the Internet and the WWW, and the publication of the first news websites, it was not long before credibility research included online news in experiments (Johnson & Kaye, 1998). Within a short period of time, a lot of credibility research focused specifically on the WWW (Abdulla et al., 2002). In turn, Human Computer Interaction (HCI) researchers quickly got involved in trying to investigate what aspects, factors, elements or interactions on a news website impacted users perception of credibility online (Alsudani & Casey, 2009; Flanagin & Metzger, 2003; Fogg & Tseng, 1999, 1999; Pirolli et al., 2009).

2.2.4 Definitions of Credibility

There are four categories of definitions of credibility in the literature. The first define it as a *perceptual variable* in the receiver of the communication. The second relate to *believability*, the ability of the communicator to convince or change opinion in the receiver. The third define it as the *product of its dimensions*, these are its constituent elements such as trustworthiness, expertise, goodwill, objectivity, bias etc. The fourth are *philosophical*. They attempt to encapsulate a complete understanding of the term, its dimensions, meaning, and effect.

2.2.4.1 Perceptual Variable

Newhagan and Nass in their work investigating the differing criteria for evaluating the credibility of newspapers and television news highlight the difficulty in defining the construct as dependent on which perspective it is perceived from. From a receiver orientated perspective they define it as: “*credibility is the degree to which an individual judges his or her perceptions to be a valid reflection of reality.*” (J. Newhagan & Nass, 1989). Metzger takes a similar approach in her work on online credibility evaluation models. She maintains that: “*definitions [of credibility] are somewhat field-specific*” (Metzger, 2007). She goes on to state “*credibility is not an objective property of a source or a piece of information; instead, it is a subjective perception on the part of the information receiver*”. In their work investigating the factors and effects of online information credibility, McKnight and Kacmar also highlight the role of perception in the individual (and believability) with their definition: “*Information credibility is defined as the extent to which one perceives the website information to be believable*” (McKnight & Kacmar, 2007).

2.2.4.2 Believability

In their seminal work in the domain, Hovland et al. while describing whether or not the audience will give credence to arguments put forward by a source stated: “*When acceptance is sought by using arguments in support of the advocated view, the perceived expertness and trustworthiness of the communicator may determine the credence given them*” (Hovland et al., 1953, p. 20). In their Interpersonal Deception Theory which focuses on deception in interactive contexts, Buller and Burgoon define credibility as: “*a constellation of judgments that message recipients make about the believability of a communicator*” (Buller & Burgoon, 1996). Fogg and Tseng investigating the elements of computer credibility take a broader stance state: “*Simply put, credibility can be defined as believability. Credible people are believable people; credible information is believable information*” (Fogg & Tseng, 1999). Johnson and Kaye in their comparison of the credibility of Internet and traditional sources also highlight the fact that credibility is a value placed by the receiver on a source of information while also highlighting the alternative way it is typically measured: “*Although credibility is typically defined in terms of worthiness of being believed, it is typically measured as a multidimensional construct* (Johnson & Kaye, 1998).

2.2.4.3 Dimensions

By defining trust and expertise as the first two dimensions of credibility, Hovland and his team set in train a movement that continues to this day. They maintain that the value given to the arguments of a source is the resulting permutation of these two dimensions. *“In any given case, the weight given a communicator’s assertions by the audience will depend upon both of these factors and the resultant value can be referred to as the ‘credibility’ of the communicator.”* (Hovland et al., 1953, p. 21). Since then, a host of credibility researchers have defined additional dimensions of the construct. Reflecting this, Gaziano and McGrath in their important work on measuring credibility state: *“Several researchers have concluded that credibility is a multidimensional concept, although the dimensions vary from study to study”* (C. Gaziano & McGrath, 1986). Self, elucidating on the increasing complexity of how the dimensions are combined to form a measure of credibility states: *“Credibility has been defined as believability, trust, perceived reliability, and dozens of other concepts and combinations of them.”* (Self, 2008, p. 435).

2.2.4.4 Philosophical

More recently researchers have been keen to put forward a more philosophical insight or holistic definition, often linking it to existing cognitive science or information science theories. Hilligoss and Rieh in their framework for credibility assessment give their perspective with: *“Credibility differs from cognitive authority and information quality, while being closely related to these two concepts”* (Hilligoss & Rieh, 2008, p. 1469). In their work “Psychological Approaches to Credibility Assessment Online”, Flanagin and Metzger state: *“In the fields of communication and psychology, credibility is traditionally defined as the believability of information, and it rests largely on the trustworthiness and expertise of the information source or message, as interpreted by the information receiver (Hovland, Janis, & Kelley, 1953; Rieh & Danielson, 2007; Tseng & Fogg, 1999)”* (Metzger & Flanagin, 2015, p. 446). They later continue: *“Thus, the fields of communication and psychology treat credibility as a perceptual variable: credibility is not an objective property of a source or a piece of information. Instead, it is a subjective perception on the part of the information receiver (Fogg & Tseng, 1999)”* (Metzger & Flanagin, 2015, p. 446).

2.2.5 Two Sides to the Coin: Defining the Construct of Credibility, and Measuring Credibility of Source, Medium, and Message

Historically credibility research has been divided into two fields. The first is predominately undertaken from within the domain of psychology and was originally concerned with developing a unifying theory of the construct. To this end, the first two dimensions of credibility, namely trustworthiness and expertise of the communicator, were defined by Hovland and his Communication and Attitude Change group in Yale (Hovland et al., 1953). Since then, researchers have continued to add new dimensions, such as bias, to the construct for source, medium, and message, in a multitude of domains and situations (Berlo et al., 1969; Cecilie Gaziano & McGrath, 1986; Markham, 1968; McCroskey, 1966; McCroskey & Jenson, 1975; McCroskey & Young, 1981; P. Meyer, 1988; Self, 2008; Whitehead, 1968). While a unifying theory has

yet to be developed, and may never be forthcoming, increasingly specific dimensions for use in specific situations continue to be defined. The second population of researchers, primarily from Communications or Information Science, rely on the increasing array of dimensions to measure the credibility of source, medium, and message. Among the wide range of credibility studies focused on source are those to measure the credibility news anchors (source as a person) (Bracken, 2006; Weibel et al., 2008). Studies investigating the credibility of a news agency (source as an organisation) (J. Newhagen & Nass, 1989). And studies investigating the credibility of e.g. a health related website (source as an information resource) (Eastin, 2006; Eysenbach, 2008). Similar sub-categories of studies exist for message and medium. While there is certainly a large overlap between both fields within the domain, an awareness of the division helps to better understand the topic.

2.2.6 The Three Meanings of ‘Dimension of Credibility’

Dimensions of credibility is a phrase with three meanings in the literature. First, dimension of credibility can refer to the constituent elements of the construct as typically defined in interpersonal communication. These include Trustworthiness, Expertise, Dynamism⁴, Goodwill, Objectivity, and Personal Attraction (Berlo et al., 1969; Giffin, 1967; Hovland et al., 1953; Markham, 1965, 1968; McCroskey & Teven, 1999; Whitehead, 1968). Figure 2-1 depicts Choi and Stvilia’s visual breakdown of the main underlying dimensions of credibility, including the discovery or definition of each of these dimensions, which highlights overlaps and commonality (W. Choi & Stvilia, 2015b). Second, dimension of credibility may refer to specific conceptualisations such as: ‘expertise of the communicator on the subject’, ‘trustworthiness in the website’, ‘accuracy of the message’, ‘fairness to both sides of the news bulletin’, ‘bias in the television network’, or ‘depth of coverage in the article’ etc. These often form the bases for measurement scales used in experiments on the subject. In most incidences, the Mean score of which is the final credibility score. Third, dimension of credibility can be used to describe the ‘dimension of credibility’ being measured, typically; source, medium, or message.

TABLE 1. Underlying dimensions of credibility in interpersonal communication.

	Trustworthiness	Expertise	Dynamism	Goodwill	Objectivity	Personal attraction
Hovland et al. (1953)	✓	✓				
Berlo et al. (1969)	✓	✓	✓			
McCroskey and Teven (1999)	✓	✓		✓		
Whitehead (1968)	✓	✓	✓		✓	
Giffin (1967)	✓	✓	✓	✓		✓

Figure 2-1 A breakdown by Choi and Stvilia of the underlying dimensions of credibility as defined in interpersonal communication and the name of the researchers who defined them (W. Choi & Stvilia, 2015b).

⁴ While almost all cite Berlo, Lemert, and Mertz as the first to define Dynamism as a core dimension of credibility, Markham highlighted its role in his Doctoral thesis “The Dimensions of Source Credibility of Television Newscasters” four years earlier and which was subsequently published (Berlo et al., 1969; Markham, 1965, 1968). Both contributions will be cited in this thesis. There is no indication that Berlo et al. was aware of Markham’s work.

Figure 2-2 depicts how a typical credibility score is derived. 1) The researcher first decides which dimension of credibility is being measured. This is typically Source, Medium, or Message. It is the focus of the investigation and informs the literature search. 2) They then select from the literature suitable dimensions of credibility (dimensions of credibility of a source in this instance), that reflect or encapsulate what they are trying to measure. Alternatively, they may put forth or argue for new dimensions of the construct depending on the situation. These are then used to identify appropriate scale item measures for e.g. trustworthiness of the communicator. 3) These are then implemented as measures via a measurement scale such as a Semantic Differential Scale or a Likert scale. In some situations, a combination of scales / measures is used to measure a single dimension of credibility such as dynamism or goodwill (Bracken, 2006). 4) A rudimentary credibility score, typically the Mean score of the individual measures, is then computed. This unweighted score does not take into account the differing importance of each of the dimensions and measures and the domain it is being measured in.

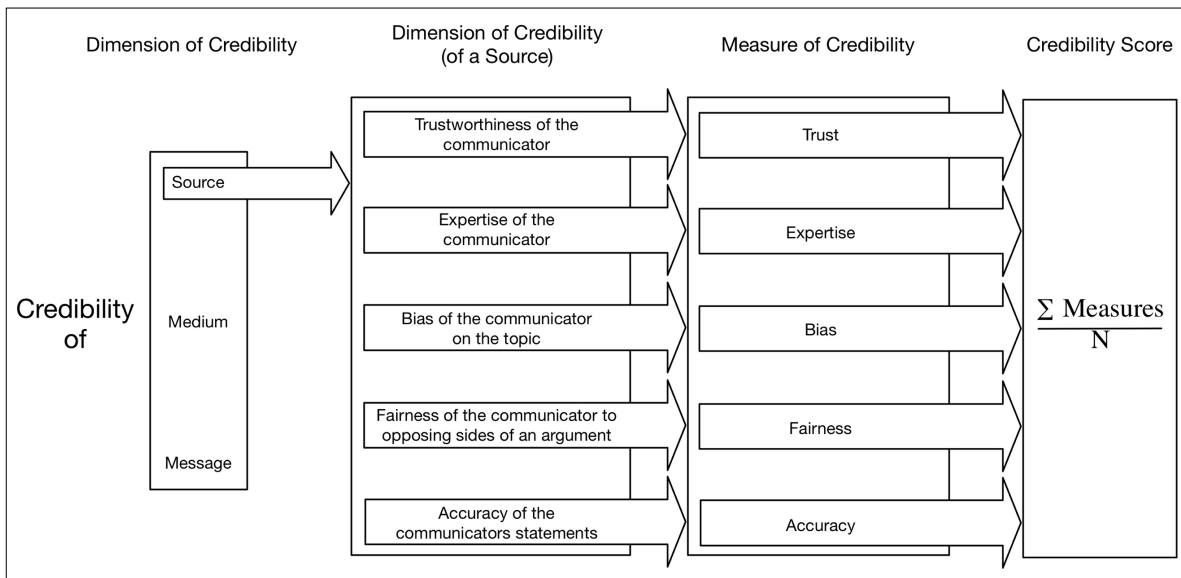


Figure 2-2 A depiction of the process of measuring credibility and how a credibility score is derived.

Six factors confuse this situation further. 1) Depending on which dimension of credibility; source, medium or message is being measured, affects which dimensions, e.g. expertise, trust etc. are typically used as measures. 2) As this chapter will demonstrate, there is no standard set or sets to measure source, medium, or message. 3) Some literature refers to either dimension as aspects of credibility or other similar terms. 4) When a dimension such as ‘expertise of the communicator on the subject’ is used in an experiment to measure the credibility of e.g. a news broadcast, it is often implemented as a single term ‘expertise’ in a Likert or Likert type scale and subsequently referred to as a ‘measure of credibility’. 5) There is some confusion over what exactly constitutes a dimension and a measure. Indeed, many studies in the domain use the terms interchangeably. In this thesis, measure of credibility is used to describe the individual measures which are implemented in some form of a measurement scale to measure the construct of

credibility in an empirical study. In some cases, this may be a single measure/scale for each dimension, in others, a researcher may choose to use several measures/scales for a single dimension. 6) Since Hovland and his team defined the first two dimensions of credibility, as trustworthiness and expertise of the communicator, researchers have been continuously defining additional, more complex, intertwined and overlapping dimensions, typically through factor analysis of survey results. This has often followed the introduction of new mediums, such as the WWW. This has created an unfortunate and confusing situation in the literature which could be partially clarified by using differentiating terms for the different dimensions of credibility.

2.2.6.1 Bias as a Core Dimension of Credibility

Bias is considered a core dimension of credibility by the majority of researchers in the domain. While Hovland and his team did not name bias as a: “*component of credibility*” of a communicator, such as expertise and trustworthiness, they were cognisant of its effect. They highlight an earlier study by Ewing when investigating the impact of: “*variations in the characteristics of a communicator with respect to expertness trustworthiness*”. Ewing’s study demonstrated that when the communicator communicates his intention to persuade, a.k.a, their bias, that there is a greater degree of change among groups with the same initial bias as the communicator. The study also demonstrates that the impact of the communicator acknowledging his bias also had a positive impact, even when the communication came from an unknown or ambiguous source. Most surprisingly, the study suggests the effect may also be true when the receivers initial bias is in direct contradiction to the communicator’s bias. They go on to state: “*these evaluations depend not only upon the content but upon the amount of conflict between the initial bias of the recipient and the avowed intention [bias] of the communicator*” (Ewing, 1942; Hovland et al., 1953, pp. 25–27). A subsequent study by Kelman and Hovland also found that persuasive messages from communicators which were perceived as negative or biased were initially less effective than those who were positive or unbiased and overall less effective long-term (Kelman & Hovland, 1953). Since then, a range of contributions, predominately from the psychology end of the credibility domain, have demonstrated the importance of bias as a dimension of credibility. Gaziano and McGrath conducted one of the largest and most formative studies in the domain with 1600 participants in phase one, and 1,002 participants in phase two, final weighted experiment sample N = 875 (Cecilie Gaziano & McGrath, 1986). The first aim of the study was to identify the dimensions of credibility. A factor analysis of the 875 participant responses revealed that 12 items from a 16-item scale grouped together as representative of the construct of credibility. The scale item (factor) “Is biased or unbiased” received a factor loading of .66 for ‘Daily Newspapers’ and a factor loading of .82 for ‘Television News’ demonstrating its importance to judgements of credibility for both mediums. Gaziano and McGrath study demonstrates, that bias is a core dimension of credibility when users were evaluating the credibility of both newspapers and television news. Similar results were found by Meyer, and by Newhagen and Nass (P. Meyer, 1988; J. Newhagen & Nass, 1989).

Bias has also been found to be a dimension of credibility when assessing news online. A study with 536 participants by Abdulla et al. found that “*online news credibility... ..was built upon trustworthiness, timeliness, and bias factors.*” (Abdulla et al., 2002). In a large study with 2,684 participants, Fogg et al. found that 11.6% of participants cited information bias as a particular concern when evaluating the credibility of information online. However, this rose to 30.2% when evaluating news online (Fogg et al., 2003). Research conducted as part of this PhD found an even high percentage. When presented with a multiple-choice question with the 25 most common dimensions of credibility, 53.58% or 217 out of 405 participants, selected bias as a core dimension of credibility, see section 5.17.5. The results are contained in Table 9-42 in the appendix.

Bias is especially important to news organisations as Fico et al. demonstrated the negative impact of increased perceptions of bias via imbalance in the story structure on the perception of credibility of the news agency behind it: “*Imbalanced story structure directly led to perceived story bias, and perceived story bias in turn led to negative evaluation of the credibility of the newspaper publishing the imbalanced story.*” (Fico et al., 2004).

2.2.6.2 *Bias as a Core Measure of Credibility*

Bias is also one of the most common measures of credibility and has been used in a broad range of empirical studies in the domain as a core measure of credibility (Abdulla et al., 2002; Chung et al., 2010; Flanagin & Metzger, 2000, 2003, 2007; K. S. Freeman & Spyridakis, 2004, 2009; Greer, 2003; Hong, 2006; Johnson & Kaye, 2002; Metzger, Flanagin, & Zwarun, 2003; Rains & Karmikel, 2009; Sundar, 1999; Sundar & Nass, 2001).

2.2.7 *Dual-Process Models Persuasion*

This section will delve the underlying theory behind almost all ten MTFS which have been proposed to explain how users form judgements of credibility online. These are the Dual-Process Models of Persuasion, namely, the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986) and the Heuristic Systematic Information Processing Model (HSM) (Chaiken, 1980). Dual-Process Models of Persuasion are one subset of a field of research in Psychology called Dual-Process Models of Social Information Processing (Chaiken & Trope, 1999; Sherman et al., 2014). The two Dual-Process Models of Persuasion are underpinned by the Controlled and Automatic Human Information Processing Models (C&AHIP) (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). There are strong overlaps between the two Dual-Process Models of Persuasion and the underpinning theory. They stipulate that when the information processing task is new, important, or when the user is engaged, that they rely on the central, systematic, or controlled, information processing routes respectively. However, when the user is unengaged, bored, or cognitively lazy, they rely on the peripheral, heuristic, or automatic, information processing routes respectively. As can be seen in Table 2-1,

eight out of ten of the MTFS reference one of these underlying theories, and many of them reference two or even all three.

2.2.7.1 Judgements of Information Via the Central Route, Systematic Strategy, and Controlled Information Process of the ELM, HSM, and C&AHIP Respectively

Although there are differences between them, in general, the central route, systematic strategy, or controlled human information process from the ELM, HSM, and C&AHIP respectively, claim that when users are motivated by task, it is important, or they are inspired, committed, or learning something new, they predominately utilise the central route, systematic approach, or controlled human information process when judging information. Thus, they make judgements based on detailed and structured evaluation of the information presented to them. This includes judgements of message factors including but is not limited to argument quality, completeness of information, balance, depth, references and citations. They will evaluate individual claims and the overall message against what they previously know, and they may cross check and evaluate individual claims and facts presented to them against third parties or other resources. This is a labour intensive, cognitively demanding and time-consuming process.

2.2.7.2 Judgements of Information Via the Peripheral Route, Heuristic Strategy, and Automatic Human Information Process of the ELM, HSM, and C&AHIP Respectively

In general, the ELM, HSM, and C&AHIP respectively hold that users will judge information via the peripheral route, heuristic strategy, and automatic human information when they are unmotivated by task, it is unimportant or inconsequential, they are uninspired, uncommitted, or they are undertaking mundane or repetitive work, or in situations of information overload, or they feel they are not able to judge the information presented to them. Thus, they will make judgements based on other factors surrounding the information. For judgements of information from a source they may rely on factors such as confidence, the level of deference or respect given to them by others, qualifications, titles and even factors such as their voice or attractiveness. For judgements of information presented in visual, aural, text, or mixed mediums, humans may evaluate a message based on, which mediums it is presented in, the organisation or source behind it, the tone of voice, attractiveness of the presenter, or its visual presentation. They will make judgements based on heuristics which are triggered by visual, aural, or information cues such as the length of a message, the use of language, and the quality of the presentation. This is a much less cognitively demanding and time-consuming method of evaluation.

2.2.7.3 Judgements of Credibility and Bias on the WWW

It is the contention of this research that users utilise the peripheral route, heuristic strategy, or automatic process to when making judgements of bias in news websites. Bias is a core dimension and measure of credibility and one which is relied on more heavily when users are assessing news online. Other researchers have also suggested that dual-process models of persuasion such as the ELM and HSM have a role in

judgements of bias by the peripheral route or heuristic strategy (Schindler et al., 2017; Turner, 2007; Yalch & Elmore-Yalch, 1984).

The dual-process models of persuasion demonstrate how this occurs when users form judgements of credibility. It may also explain how it happens when users form judgements of bias, which is a core dimension and measure of credibility when judging news online.

2.2.7.3.1 *The Elaboration Likelihood Model (ELM)*

Figure 2-3 is a reproduction of Petty and Cacioppo's ELM from their 1986 paper and based on their earlier work (Petty, 1977; Petty & Cacioppo, 1978, 1981, 1986). The ELM is referenced in seven of the ten MTFS highlighted in Table 2-1. It proposes the central and peripheral routes of information processing. Petty and Cacioppo maintain that when users are motivated, either by personal relevance, responsibility, or their ability to process issue related information is high, they use the central route. This involves cognitive elaboration, and judgements of information and argument quality. However, when their motivation is low, due to tiredness, information overload, or the unimportance of the task etc. and there are peripheral cues present, they will likely take the peripheral route to judgements of credibility. As Gass and Seiter frankly describe in their book 'Persuasion: Social Influence and Compliance Gaining': "*to put it bluntly, low-involved receivers are cognitively lazy*" (Gass & Seiter, 2015). The peripheral route is much less cognitively taxing than the central route.

As Sundar and others propose, websites are rife with technical and visual cues by which users can use as a peripheral route to form judgements of credibility (Sundar, 2008). These range from the obvious, such as reputable domain names and recognisable brands, to the less obvious such as names and profiles of people behind the website or having privacy policies and terms and conditions. As can be seen in the ELM shown in Figure 2-3, a failing of motivation or ability to process will encourage users to make assessments via the peripheral route. Therefore, they are more likely to make credibility judgements not based on the information presented to them, or on the validity or truthfulness of the arguments in the message, but rather by the peripheral route which predominantly relies on visual cues. These may be contained in individual technical features, or the combination of them in the website or webpage, or in the visual quality based on characteristics such as professionalism, gaudiness, or the overall look and feel of the website or webpage. This includes the combination and balance of colours, layout, structure, accompanying graphics, video, or imagery etc. They may also make judgements of technical competence or authority, which may contribute to perceived credibility, based on the quality of the technical implementation. Wathen and Burkell believe that the combination of cognitive authority and technical authority create institutional authority which is

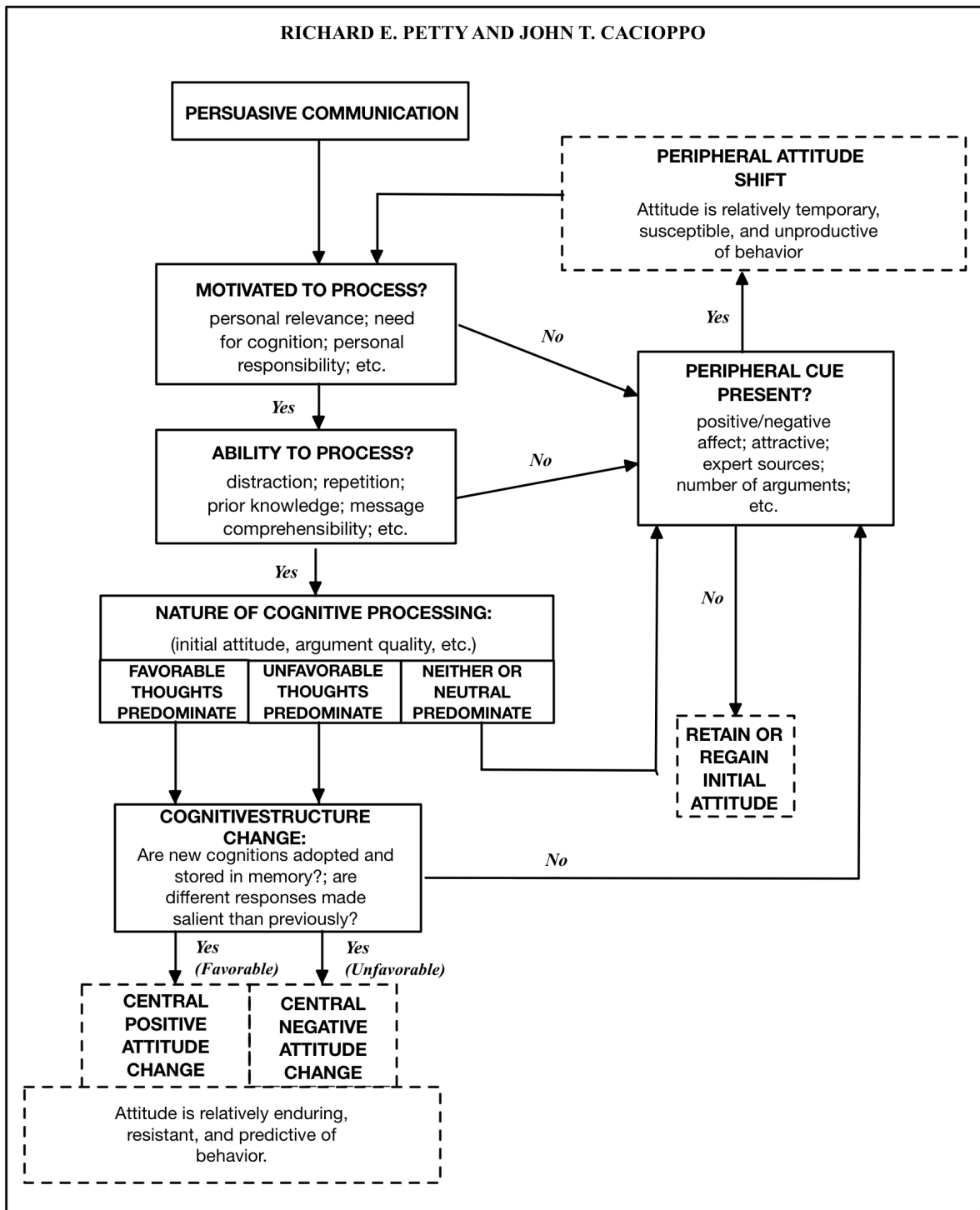


Figure 2-3 The Elaboration Likelihood Model of Persuasion (ELM), adapted from (Petty, 1977; Petty & Cacioppo, 1978, 1981, 1986).

important in judgements of credibility (Wathen & Burkell, 2002). Users may even make visual judgements of the content such as length, paragraph structure, loudness or gaudiness of headings, use of inline links, and bullet points, as users were found to do in one recent study on credibility (W. S. Jung et al., 2017).

As can be seen in Table 2-1, the ELM provides much of the underpinning theory to seven of the ten MTFS. If anything, it seems more suited to describing routes for effective persuasive communication on the WWW than in the traditional mediums of its time due to the abundance of available visual cues and the amount of time users engage in casual browsing. Hilligoss and Rieh also provide a detailed explanation of the ELM and how it pertains to web browsing (Hilligoss & Rieh, 2008).

2.2.7.3.2 *The Heuristic Systematic Model (HSM)*

The Heuristic Systematic Model of Information Processing (HSM) see Figure 2-4⁵, proposed by Chaiken, is another Dual-Process Models of Persuasion (Chaiken, 1980). Its systematic and heuristic processing strategies are analogous to the central and peripheral routes in the ELM. It is referenced in four of the ten MTFS presented in Table 2-1 and was the main contribution of Chaiken's Doctoral thesis (Chaiken, 1977). The experiments which contributed to the development of the HSM demonstrated that those with a high involvement used a systematic processing strategy of the information, while those with a low involvement used a heuristic strategy of information processing via peripheral cues (Chaiken, 1980). As none appears to exist in the literature, a visual representation of the HSM has been created based on its description in the original paper, see Figure 2-4. The HSM proposes that users employ two different strategies, systematic and heuristic, to process information. The systematic strategy is high involvement. Users process persuasive argumentation in the content or message in a thoughtful, analytical, and deliberate manner. However, it is time consuming and requires high cognitive effort. It is more likely to be employed when reliability concerns are high, or when accuracy or veracity are important. In contrast, the heuristic processing strategy is low involvement and may be employed when judgements are inconsequential, or the topic or process is unimportant. Users rely on simple or general decision rules or cognitive heuristics based on previous experience to evaluate information. They also rely more on non-content cues or source cues.

The proportion of user browsing that is low involvement and high involvement fluctuates for different users, however it is undeniable that for almost everyone, at least some browsing is low involvement, unfocused, spent on mundane tasks, or without objectives. While for others, it is a significantly larger proportion of time. Therefore, per the HSM, almost all users engage in heuristic browsing strategies upon occasion and may make judgements of credibility based on source cues or other non-content cues, such as the presentation of the content or its underlying technical affordances.

⁵ Despite the HSM being one of the most highly cited Dual-Process Models of Persuasion, second only to the ELM, it appears that it has never been realised in visual form. Thus, its effectiveness for quickly conveying information, and its popularity, is reduced compared to the ELM. This representation of the HSM has been included in a journal paper which is currently being prepared for submission.

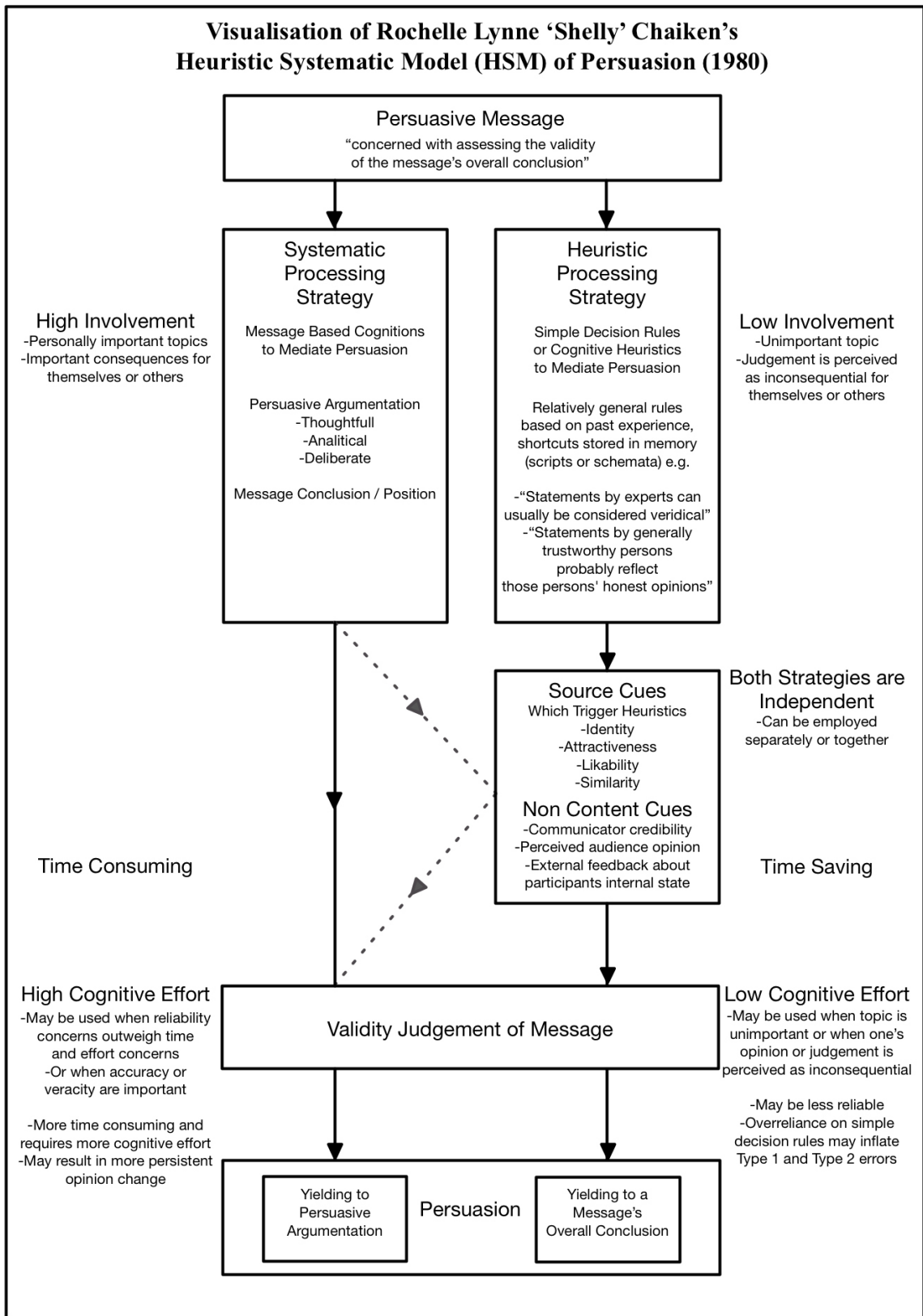


Figure 2-4 Visual representation of Chaiken’s Heuristic Systematic Model of Information Processing (HSM) which is described, but not visualised, in the paper “Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion” (Chaiken, 1980).

While the HSM has its detractors, mainly due to its perceived similarity to the ELM, it sufficiently differentiates itself enough from the ELM to be a notable contribution in its own right. This is most evident in its focus on source cues and other non-content cues, which both independent strategies may avail of. It is also notable for highlighting the independence of the strategies, while also highlighting the fact that both may be employed at the same time.

2.2.7.3.3 *Controlled and Automatic Human Information Processes (C&AHIP)*

Despite the attention given to the ELM, and to a lesser degree the HSM, the seminal underpinning work in this area is Schneider and Shiffrin's Controlled and Automatic Human Information Processes (C&AHIP) (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). It is referenced twice by the ten MTFS presented in Table 2-1. Their work and extensive experimentation resulted in a general theory and framework of information processing. It provides the underpinning theory for the ELM and HSM. Part One of their highly cited and in-depth paper proposes a two process theory of detection, search, and attention, and details a series of studies which traces these to controlled information processing and automatic information processing (Schneider & Shiffrin, 1977). They provide evidence demonstrating that users tend to employ controlled processing when the task is new, important, or when long term learning is required. Automatic processing is triggered by appropriate inputs from previous experience allowing the participant to detect important elements with low cognitive input. Part Two of their paper further tests the two processes and puts forward a general framework and theory for human information processing (Shiffrin & Schneider, 1977).

Each of the three theories are different arcs of perspective on the same problem. When users are not motivated by task or importance, they will adopt less cognitively demanding peripheral, heuristic, or automatic information processing strategies that rely on visual cues surrounding the content or message. These Dual-Process Models of Persuasion and the C&AHIP underpin most the ten MTFS detailed in section 2.2.8 and highlighted in Table 2-1.

One of the main suppositions behind this research, is that due to the constant prevalence and availability of news on the WWW, users have heavily adopted such peripheral, heuristic, or automatic methods of evaluating the news that they consume online. These may allow them to quickly judge the acceptance of the information they are consuming. Combined with the fact that bias is a core dimension and measure of credibility, one of the most important considerations when consuming news, and given the fact that website design is already known to impact perceived credibility, it is possible that there are cues or triggers in a news website's design which may impact the perception of bias.

2.2.8 Human-Centric Models, Theories, Frameworks, and Schematics (MTFS) for Judging or Measuring Credibility Online

An exhaustive review of the domain identified ten Models, Theories, Frameworks, and Schematics (MTFS) to explain how users evaluate the credibility of information on the WWW. Each of these takes a human-centric approach to explain how users assess source, medium, and message credibility online. The review found no empirical frameworks for measuring credibility. The following provide an overview of the MTFS in chronological order.

Fogg and Tseng's 'Three Models of Credibility Evaluation' was the earliest discovered contribution in the domain (Fogg & Tseng, 1999). While this relates to 'computing products' rather than online credibility, it is the forerunner in the domain and provides a valuable contribution. Fogg and Tseng propose three models, binary, threshold, and spectral to explain how users evaluate credibility of computing products. They propose that as users' interest, ability, familiarity with the subject matter, and ability to compare multiple sources increase, that they are more likely to rely on spectral evaluation rather than threshold or the most basic, binary evaluation. This paper was based on a synthesis of existing research and theory rather than experimental evaluation. The paper was designed as a conceptual framework, to define key terminology and to promote further research in the field.

Fritch and Cromwell's 'Model for Ascribing Cognitive Authority to Internet Information' is the earliest attempt specifically relating to how individuals assess the credibility of information online (Fritch & Cromwell, 2001). Their iterative model maintains that users form judgments of the document, author, institution, and affiliations, and that these combine to form an overall credibility judgment. They argue that by ascribing cognitive authority, first put forward by Wilson (Wilson, 1983), to information found online that information seekers are able to form a judgment.

Rieh and Belkin's qualitative studies observing scholars search behaviour on the web formed the basis of Rieh's 'Model of Judgment of Information Quality and Cognitive Authority' (Rieh, 2002; Rieh & Belkin, 2000, 1998). The model drew heavily on existing theory and models in the wider fields of human judgment and decision making in Information Retrieval. It focuses on information quality and cognitive authority as two of the main factors apparent in the existing literature in assessing information online. The model reflects the fact that scholars in the contributory studies were found to have two distinct types of judgment: predictive judgment and evaluative judgment, which were first put forward in Hogarth's work on judgment and choice (Hogarth, 1987). Rieh's model differs from other MTFS of credibility assessment as it focuses almost completely on the factors influencing the act of judging information. This includes the characteristics of the information and its source, the user's knowledge, and the task, and how they impact on perceived information quality and cognitive authority.

Wathen and Burkell's contribution is their proposed 'Model for how Users Judge the Credibility of On-Line Information' (Wathen & Burkell, 2002). Their paper provides a synthesis of the factors influencing five aspects of credibility from the literature and highlights earlier work by Olsen and others that divide these factors into those related to 'cognitive authority' and 'technical authority'. In the absence of human sources, Wathen and Burkell, citing Olsen, maintain that users assign cognitive authority to the author of the message on a website. The combination of cognitive and technical authority creates 'institutional quality'. In support of their position, they highlight the earlier work of Rieh and Belkin who found that academics seeking information online look for surface credibility markers such as academic or governmental URLs, and/or statements of the authors credentials and institutional affiliations, which substitute traditional means of establishing credibility (Rieh & Belkin, 2000, 1998). Wathen and Burkell recognised that there are a host of influences involved in evaluating the credibility of information online and the exact makeup and influence of these factors change depending on the user, context, and urgency (Fogg et al., 2001; Rieh & Belkin, 1998). Consequently, they propose a much more complex, multistage iterative model of credibility assessment than previous efforts. The first stage, Evaluation of Surface Credibility, involves assessments of appearance or presentation, usability and interface design, and the overall organisation of the information. The second stage, Evaluation of Message Credibility, involves the assessments of source through perceived expertise, credentials etc. and assessments of the message in terms of content, relevance, currency, accuracy and tailoring. The final stage, Content Evaluation, involves the users making an evaluation of the actual textual or alternative content. At any stage the user may judge an information resource or message as not being credible and seek an alternative. The model proposes that the user evaluates each of these stages by asking themselves different questions such as: "*Does this site look professional?*". The majority of subsequent contributions have adopted this principle in one form or other.

Fogg's second and more relevant contribution is 'Prominence-Interpretation Theory' (PIT) (Fogg, 2003). The theory proposes that visitors notice individual elements of a website or webpage and then evaluate each to construct an overall credibility score. The theory is based on a reflective analysis of the process a user goes through when they make a credibility assessment. It maintains that two things happen during this process, firstly the user notices something, Prominence, and secondly the user makes a judgment, Interpretation. This happens multiple times in an iterative process while the user is browsing a website. The theory explains that at least five factors affect Prominence, which is an elements likelihood of being noticed during credibility evaluation. They include; involvement, topic, task, experience, and individual differences. And that at least three factors affect Interpretation, which is the value or meaning that the users assigns to the element and whether it is good or bad. They include; assumptions, skill/knowledge, and context. Fogg's earlier contribution, along with Tseng detailed above, is one of the formative studies that gave rise to PIT. It proposed three types of credibility evaluation, namely, binary, threshold, and spectral evaluation, and that depending on the situation, users adopt one of the three strategies depending on the

task and the information or feature (Fogg & Tseng, 1999). However, PIT does not appear to adopt Fogg and Tseng's earlier three strategy theoretical proposal.

Metzger's contribution is the 'Dual Processing Model of Credibility Assessment' (Metzger, 2007). Her work was heavily influenced by the Dual-Process Models of Persuasion in cognitive and social psychology, namely Petty and Cacioppo's Elaboration Likelihood Model, (ELM), Chaiken's Heuristic-Systematic Information Processing Model (HSM) and by the underpinning theory provided by Shiffrin and Schneider's general theory resulting from their Controlled and Automatic Human Information Processes (C&AHIP (CHIP & AHIP)), (Chaiken, 1980; Petty & Cacioppo, 1986; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Metzger proposed that like the aforementioned dual models, motivation and ability plays a central role in whether information seekers online process information by the central route, systematic approach, or controlled processes (see section 2.2.7.1), or the peripheral route, heuristic strategy, or automatic processes (see section 2.2.7.2). Metzger's abstract model involves three phases; exposure, evaluation, and judgment. Metzger claims that her Dual Processing Model of Credibility Assessment might be useful in predicting whether users will take the central route, systematic approach, or controlled model; or the peripheral route, heuristic strategy, or automatic model, of the ELM, HSM, or the C&AHIP when processing information online.

Sundar's 'MAIN Model' (Modality, Agency, Interactivity and Navigability) differs from earlier work in the domain by focusing on the effects of technology on the user (Sundar, 2008). Sundar reasons that although credibility is typically considered by assessing the source of the information, this is not always possible due to the multiple layers that information is filtered through on the WWW. Citing the long established and heavily cited ELM, Sundar argues that social psychologists have long claimed that there are peripheral cues in persuasive messages in traditional media and that viewers are able to use these peripheral cues as an alternative peripheral route to arrive at a conclusion (Chaiken, 1980; Petty & Cacioppo, 1986). This is as opposed to the central route, systematic approach, or controlled model of assessing the message. Sundar maintains that for online information seekers, information is processed in a similar fashion, and that there is a greater range of peripheral routes than in traditional media. Viewers can systematically process the text of a message to isolate relevant information to form a judgment while at the same time evaluating its presentation heuristically in terms such as its length, structure, readability and presentation. Sundar asserts that the fundamental source of these heuristics is the differing underlying affordances in the technology behind the medium of communication. Citing the seminal work of Schneider and Shiffrin's Controlled and Automatic Human Information Processing Models (C&AHIP), Sundar also highlights how users may unknowingly be using heuristics to form such judgements (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). He maintains that there are two means by which affordances convey

credibility cues to the user. First, by their sheer presence, such as the interactivity affordance⁶, and second by assembling information, related to the message or the topic such as how often a video has been watched. Sundar maintains that these peripheral cues can emit *information scent* that is pertinent to the user (Sundar et al., 2007). These surrounding elements to the central content, or *information scent*, allow the consumer to make quick judgements of the quality of the information they are consuming. The importance of each affordance and their resulting cues depends on several factors and the heuristics they invoke may have a positive or negative effect depending on elements such as salience or context.

Hilligoss and Rieh's 'Unifying Framework of Credibility Assessment' identifies three distinct levels of credibility judgment: construct, heuristic, and interaction (Hilligoss & Rieh, 2008). The major contribution of this framework is the addition of a construct level to the evaluation process. It describes how the information seeker defines credibility, what characteristics of source, elements of the message, or aspects of the medium they consider salient. It also considers time and task. This enables them to iteratively combine judgments of credibility. The heuristic level defines general rules of thumb to make credibility judgements. These may be media related, source related, endorsement related, or aesthetic based. The interaction level pertains to credibility judgements based on cues. These include content cues, peripheral source cues, and peripheral information object cues. Compared to previous attempts to explain how users make judgements of credibility, Hilligoss and Rieh's framework is more of an abstract consideration of the factors involved. The other key contributions of the framework are highlighting the strong influence of context surrounding assessments of credibility and showing that information's seekers regularly check and compare several information sources at the same time.

In their paper 'Revised 3S-Model of Credibility Evaluation', Lucassen et al. propose three strategies of feature evaluation; semantic, surface, or source, which the information seeker can use to arrive at a judgement of trust (Lucassen et al., 2013). The choice of strategy is based on user characteristics; domain expertise, information skills, or source expertise, and one or more strategies can be followed concurrently. What differentiates this from other attempts to explain how users form judgements of credibility is its almost complete focus on trust. They maintain that credibility is the aspect of information or the communicator which the receiver puts their trust in.

Metzger and Flanagin's 'Schematic Representation of Potential Factors Relevant to Credibility Evaluation of Online Information' is an attempt to better understand psychological approaches of users to judgements of credibility online (Metzger & Flanagin, 2015). There are two contributions of note from this research. First, from their review of the domain and analysis of existing models and frameworks, they developed a

⁶ Sundar maintains that: "Each technology brings with it a set of affordances or capabilities that can shape the nature of content in a given medium. In addition to dictating content, these affordances also determine the way the content is typically presented via the medium and receivers' states of mind while using it. For example, the affordance of interactivity on a Web site suggests 'action possibilities' such as clicking on hyperlinks or typing in a chatroom. These possibilities suggest openness of information access and the participatory nature of the forum, among other things" (Sundar, 2008).

schematic of the various cues that may trigger heuristic judgments. These are the means by which users make quick judgements of the credibility of the information presented to them. It is the contention of this research that users use a similar methodology when judging bias in news online. Second, is their proposal for how users might evaluate information from social media or in an online social context. While not a model, theory, or framework, explaining how users make credibility judgements online, its comprehensive categorisation of cues involved in credibility evaluation decisions warrants its inclusion.

The latest contribution in this domain is Choi and Stvilia's "A New Framework for Web Credibility Assessment". This framework, presented in a poster at iConference 2015, has been devised from a literature analysis that conceptualises the relationships between the key dimensions of credibility, their related measures, and the dimension⁷ of credibility being measured, source, message, or medium (W. Choi & Stvilia, 2015a). The background to this research, including Choi and Stvilia's alternative analysis and breakdown of six of the frameworks highlighted here, is available in their earlier review of the domain (W. Choi & Stvilia, 2015b). Their review uses four facets; context, user characteristics, operationalization, and process, to highlight the commonalities and differences between the different frameworks. It should be noted though that this work has yet to research maturity or to be published in long form.

The MTFS reviewed heretofore are the major works in the literature that attempt to explain how users form credibility judgments online. Table 2-1 presents a summary in chronological order from left to right. The contribution of Choi and Stvilia is not included due to insufficient published literature available on it. Table 2-1 also shows which dual-process model of persuasion route or information processing strategy each relies on and whether they are based on empirical investigation, existing theory, or both. The key advancement of each is then highlighted followed by the major influences from the underlying theory. Despite this exhaustive review, no empirical framework was identified. This clearly shows the gap for a formal, open research framework, which can be adapted to individual purposes.

It is worth noting that only Fritch and Cromwell's and Rieh's models focus on the central route, systematic approach, or controlled model from the ELM, HSM, or the C&AHIP. The later models and theories increasingly focus on the heuristic, peripheral routes or the automatic model. This is due to the knowledge that users rely on cues surrounding the content or the message contained therein for judgments of credibility. Each MTFS forms a different arc of understanding, centred on the same problem and has added to the body of knowledge. Nevertheless, singularly and combined, they demonstrate that there is still a gap in the canon as to how different features and characteristics of a website's design, reflected in a webpage's aesthetic, impact on the perception of credibility and each of its individual dimensions and measures.

⁷ Choi and Stvilia interchangeably use the terms such as "objects of credibility assessment", "objects of assessment", "objects under investigation" etc. to describe the dimension of credibility being measured, source, medium and message. However, this has not been adopted in the wider domain (W. Choi & Stvilia, 2015a, 2015b)

Table 2-1 Ten human-centric Models, Theories, Frameworks, and Schematics (MTFS) explaining how users form judgements of credibility on the WWW.

Name	Three Models of Credibility Evaluation	Model for Ascribing Cognitive Authority to Internet Information	Model of Judgment of Information Quality and Cognitive Authority	Model for how Users Judge the Credibility of On-Line Information	Prominence Interpretation Theory (PIT)	Dual Processing Model of Website Credibility	MAIN Model	Unifying Framework of Credibility Assessment	3S-Model of Credibility Evaluation	Schematic Representation of Potential Factors Relevant to Credibility Evaluation of Online Information
Creators	(Fogg & Tseng, 1999)	(Fritch & Cromwell, 2001)	(Rieh, 2002)	(Wathen & Burkell, 2002)	(Fogg, 2003)	(Metzger, 2007)	(Sundar, 2008)	(Hillgoss and Rieh, 2008)	(Lucassen, 2013)	Metzger & Flanagin, 2015)
Dual-Process Model of Persuasion Route or Information Processing Strategy	Central / Systematic And Peripheral / Heuristic	Central / Systematic	Almost totally Central / Systematic Information Quality and Cognitive Authority are systematic	Stage 1 Peripheral / Heuristic Stage 2 Central / Systematic	Almost totally Peripheral / Heuristic	Peripheral / Heuristic Central / Systematic	Almost totally Peripheral / Heuristic	Heuristic Layer Peripheral / Heuristic Interaction Layer Central / Systematic	Unfamiliar with topic Peripheral / Heuristic Familiar with topic Central / Systematic	Peripheral / Heuristic Central / Systematic
Based on Experiment or Theory	Theory	Theory	Theory based on Experiment	Theory	Theory based on previous Experiment	Theory based on previous Experiment	Theory	Theory based on Experiment	Theory based on Experiment	Theory
Advancement	First effort in the domain to model how credibility judgements are made. Defined key terms	Introduces the concept of multiple concurrent filters or information processing strategies which work iteratively and can form judgements of cognitive authority	Introduces Predictive Judgement and Evaluative Judgement	Two stage Iterative model. It proposes that users are asking themselves questions, the failing of which will result in them leaving the website	Iterative process where users notice something on a website which either increases or decreases perceived credibility	Three phases: Exposure Evaluation Judgement. Type of evaluation heavily influenced by motivation as per ELM	Underlying technologies have Affordances, each of which act as a repository of cues, some of which will aid judgements of credibility.	Adds a construct layer, what does the assessor conceptualise or define credibility as First to consider context	Discovered that those familiar with a topic tend to focus on systematic evaluation while those unfamiliar focused on heuristic. First to consider user generated content.	First to consider social media. Identified three sets of cues: Website or Source, Message and Author cues, which combined with the receiver's characteristics determine credibility
Influenced by	ELM Petty and Cacioppo 1981, 1981 Chaiken 1979	None of them	None of them	ELM Petty and Cacioppo 1986, 1990	ELM Petty and Cacioppo 1986	HSM Chaiken 1980, Chaiken & Trope 1999 ELM Petty and Cacioppo 1981 C&AHIP Shiffrin and Schneider 1977	HSM Chaiken 1980, 1987 Chen and Chaiken 1999 ELM Petty and Cacioppo 1986 C&AHIP Shiffrin and Schneider 1977	ELM Petty and Cacioppo 1981, 1986	HSM Chaiken 1980,	HSM Chaiken 1980, 1987, Chaiken & Trope 1999 ELM Petty and Cacioppo 1981

2.2.9 *Connecting the Underlying Theory of the ELM, HSM, and the C&AHIP, to the MTFS of Credibility Judgement, and the Visual Presentation of News Webpages*

Like the ten MTFS explaining how users form judgements of credibility on the WWW that were highlighted in Table 2-1, the ELM, HSM, and C&AHIP detailed above in section 2.2.8, are different arcs of perspective on the same problem. Broadly speaking, each demonstrates how users may employ one of two strategies while browsing or consuming information. When users are motivated by task, it is important, or they are inspired, committed, or learning something new, they predominately utilise the central route, systematic approach, or controlled human information process of credibility evaluation from the ELM, HSM, and C&AHIP respectively. When users are unmotivated by task, or it is unimportant, or they are uninspired, uncommitted, or they are undertaking mundane or repetitive work, or they are experiencing information overload, or they feel they are not able to judge the information before them, they will adopt the less cognitively demanding peripheral route, heuristic strategy, or automatic human information process to evaluate credibility from the ELM, HSM, or C&AHIP. Eight out of ten of the MTFS directly reference and heavily rely on either the ELM, HSM, or C&AHIP as their underlying theory, and many of them reference two or even all three.

2.2.9.1 *Connecting the MTFS of Credibility Judgement to Individual Features and Characteristics of a News Website's Design, Reflected in a news Webpage's Aesthetic*

In the earliest model of credibility evaluation by Fogg and Tseng, they note the importance of the overall visual presentation of a website to judgements of credibility with: *“a web page may appear credible just because of its visual design”* (Fogg & Tseng, 1999). They also highlight the ELM and note that designers need to account for users with a low involvement or who lack ability or motivation if judgements of credibility are important.

Fritch and Cromwell's, and to a lesser extent Rieh's models focus on the central, systematic, or controlled information process for judgements of credibility. Yet, Fritch and Cromwell's still note the importance of format and presentation, while Rieh's notes the importance of organisation/structure, presentation, graphics, and functionality. The format, presentation, organisation/structure, and graphics are all directly derived from the overall design and its characteristics, while functionality is typically provided by individual features.

Wathen and Burkell's contribution maintains that users first evaluate credibility via: *“surface characteristics”* before they evaluate the message (Wathen & Burkell, 2002). These surface characteristics include colours, graphics, font size, clutter, no obvious errors or typos, and attention to detail. All the while the user is asking themselves questions such as: *“Does the site look professional?”* It is only once these characteristics are judged acceptable that the user then moves on to evaluate the message.

Fogg's Prominence Interpretation Theory (PIT) focuses heavily on judgements of credibility based on the visual presentation. The theory maintains that two things happen in such judgements:

- “1. The user notices something (Prominence), and*
- 2. The user makes a judgement about it (Interpretation).”*

This is heavily focused on individual features of a website or webpage such as a privacy policy, graphics, links, advertising, and navigation. This theory was one of the reasons the first experiment presented in chapter 4 focused on the impact of individual features of a news website on the perception of bias.

The format of Metzger's 'Dual Processing Model of Web Site Credibility Assessment' draws heavily from the ELM and to a lesser degree the HSM and C&AHIP (Metzger, 2007). She notes that: *“Users who are less motivated to find high-quality, credible information ... may not assess credibility at all or do so by simply considering such simple characteristics or heuristics as a Web page's design and graphics.”* Metzger's pointing out that users may make judgements of the credibility of the information contained in a website (at least initially) based on characteristics of the design, which are then judged heuristically, was a contributing reason why in this research we decided to focus on the impact of characteristics of a news website's design, reflected in a webpage's aesthetic, on the perception of bias in the news articles contained therein. This focus is reflected in the experimentation presented in chapter 5 and chapter 6.

Sundar's MAIN model was perhaps the strongest influence on the decision to investigate the impact of features and to a lesser extent the characteristics of a news website's design, reflected in a news webpage's aesthetic, on the perception of bias (Sundar, 2008). Sundar continually references the importance of individual surface level features as visual cues by which users may make heuristic judgements about the credibility of a website. He also highlights the importance of surface level visual characteristics of a website's design which are: *“associated with powerful first impressions of Web site credibility.”*

2.2.9.2 Examples of Individual Features and/or Characteristics by which a User May Make Judgements of Credibility via the Peripheral Route, Heuristic Strategy, or the Automatic Human Information Process

Many of the ten MTFS highlighted earlier describe how users, taking a peripheral route or heuristic strategy, or employing automatic information processing, may use such features or others to determine the credibility of the resource. As Sundar points out, these features (and many other smaller individual features) may act as a repository of cues, as to the underlying technical affordances, which trigger heuristics which the participant uses to arrive at a credibility evaluation (Sundar, 2008). One example of a heuristic put

forward by Sundar, is the length is strength heuristic⁸, whereby a user may perceive long articles as being more credible than shorter articles without having read them. Other examples include the audio-visual heuristic, which may be triggered by visual content such as the presence of pictures or video, or the interaction heuristic which may be triggered by the presence of a comment section or other means of interacting with the information originator. There is also evidence of a numbers heuristic which may be triggered by the presence of numbers or data in the text or the presence of data driven technical features such as live share price features (Yalch & Elmore-Yalch, 1984). User may also evaluate the credibility of a website or webpage by whichever visual characteristics they perceive.

Each of the ten MTFS detailed in section 2.2.8 and highlighted in Table 2-1, put toward evolving, yet similar methods, by which users evaluate credibility online. Many of the most recent have a strong focus on credibility evaluations via heuristic strategies, the peripheral route, or automatic information processes, which users judge via features or characteristics of the overall news website's design, reflected in a news webpage's aesthetic. This is opposed to central or systematic routes, or controlled information processing strategies, which focus on judgements of cognitive authority or information quality.

As bias is a core dimension and measure of credibility, especially when users are judging information online, one of the key motivations of this thesis is to ascertain whether or not the perception of bias in online news articles was similarly impacted by its presentation and whether or not it could be predictively impacted.

2.2.10 News Website Design, News Webpage Aesthetics, and the Production of News Online

A news webpage's aesthetics is descended from a news website's design, which in turn is derived from a news organisation's overall branding and style. However, unlike static designed products, websites are regularly updated with new content, and the design may also be updated or changed as required. Compared to most websites, news websites are in a state of virtual flux. They can be updated thousands of times per day by a team of journalists, editors, designers and developers. These updates can range from the creation of a whole new section of webpages dedicated to an event of global importance with live updating and video feeds, to simple corrections to the text of an article to remedy a misspelled name. Most of the larger international news websites use Content Management Systems (CMS) costing millions of dollars, see footnote 29 in section 4.1. These are pieces of software designed to manage the production and distribution of news. When a journalist submits an article, it is passed along to their sub-editors for content checking, legal compliance, ethics etc. The sub-editors and editors may amend the language or headlines of an article, commission graphics or attach pictures. When they decide to publish the article, the content management

⁸ "For instance, a long message carries with it the length cue, which at a glance can trigger the 'length implies strength' heuristic, leading to the conclusion that the message is strong—a conclusion drawn without taking into consideration the actual content of the message." (Sundar, 2008).

system is responsible for pushing this to the various distribution channels including print, email and ezine, smartphone apps including push notifications, and news websites. Many of these channels may have their own editors or sub-editors who are responsible for layout or positioning. Depending on the medium, they may make further changes. This may include shortening an article to fit available column inches in a print newspaper or commissioning interactive graphics and features in an online article. For large multiformat news agencies such as the BBC, the multiple mediums and channels add even more complexity.

After an article has been signed off by an editor or sub-editor as being ready for publication, which depending on a number of factors may include consultation with the journalist who authored it, the journalist and even the editor and sub-editor may have little say in how it is finally presented. In print newspapers articles are subject to last minute changes based on factors such as space, position and their importance versus other news articles importance. In online news, where small differences in design or presentation can have a big effect, the headlines, images etc. of news articles on a homepage may be tweaked in an attempt to increase click-rates or the amount of time a user spends on a news website, though it must be noted that in most reputable news agencies the main content of the article is unlikely to change. In fact, as shown by Lowrey, many online editors and designers maintain that they have a large say in the content of article headlines (Lowrey, 2002). This is discussed further in section 9.5.2 in the appendix.

Bell from the Tow Center for Digital Journalism at Columbia University publishing in the Columbia Journalism Review also points to a “culture clash” between journalists and editors and the technical developers and designers who have increasing levels of control over presentation 9.5.2. The designers, developers and online editors are constantly implementing small changes to the design and presentation, relying on tools such as live A/B testing and personalised recommendations to try and maximise the click-rate and time consumers spend on articles and on the website in general. These are important to generate increased advertising revenue and to attempt to fulfil consumers news needs so that they spend more time on their news website and feel less desire to visit alternative news sources. Bell is not alone in her concerns. Ananny and Crawford at the Nieman Lab for journalism at Harvard have also highlighted the increasingly powerful role that designers have in communicating news to consumers. They argue that “*press ethics are intertwined with platform design ethics, and press freedom is shared with software designers*” (Ananny & Crawford, 2014). They coined the term “*liminal press*” to describe the people and systems that exist outside and alongside online news organisations that enable the distribution of news (Ananny & Crawford, 2015).

2.2.11 The Impact of Features and Characteristics of a Website’s Design, Reflected in a Webpage’s Aesthetic, on the Perception of Credibility

Website features are the components of a website that may include or combine text, images and interactive elements, which are used to convey information or provide a service. They are distinct from message features such as argument quality, readability, expert quotations, and rigor (O’Keefe, 2015). Through their

inclusion, omission or distortion, they may increase or decrease the perceived level of credibility among different users.

Characteristics of a website's design are the overarching themes and considerations behind a website's look and feel. Part of the overall design philosophy of a website, they reflect the attitude and values of the organisation behind the website is trying to convey to the user. Individual examples of characteristics include modern, minimalist, quality, traditional, conservative, busy and/or active. These may be transmitted through coordinated combinations of branding, colour, layout, the number of elements on a page, the inclusion of advertising and its gaudiness, professionalism, and the content of images and even the style of fonts. The most obvious and common manifestation of differing design influences in the news domain, is the dichotomy in the visual presentation between tabloid and traditional news websites.

When compared to traditional mediums, the line between content and design in news websites has become increasingly blurred. Although the majority of existing human centric credibility judgement frameworks in the domain highlight the role that surface features of a website's design play when the credibility of a website is being assessed, the work of Sundar is especially important (Sundar, 2008). In his MAIN model, he argues that the technologies behind the medium have their own affordances that bring their own set of cues, dictate how content is displayed, and determine the user's state of mind while using it. Sundar states that there are at least two ways digital technologies can convey cues which are important to assessments of credibility. The first is by their sheer presence, such as interactivity which may trigger the 'dialogue' or 'contingency' cues. The second method by which the underlying digital technology affordances can convey cues are auto-generated indicators of popularity or network size. These emit 'information scent' which Sundar claims is helpful in making judgements of credibility. Subsequently, Giudice in her investigation of social feedback on credibility demonstrated that positive or negative social feedback can influence the perception of credibility (Giudice, 2010).

Sundar contends that visual cues transmitted by features or the structure of the website are likely to be particularly important, especially to younger users who are more au-fait with quickly judging technology (Sundar, 2008, p. 76). He argues that rather than trying to persuade users to assess information via the central route, systematic approach, or controlled processes, it would be more beneficial to gain an understanding of how they assess it heuristically. In a study with 2,684 participants, in which open-ended comments were collected relating to how participants judge the credibility of information online, 46.1% of participants comments related to *Design Look*, while an additional 28.5% related to *Information Design/Structure* when making credibility evaluations (Fogg et al., 2003). Lending further weight to the argument that design and presentation impact on perceived credibility, an experiment by Robins and Holmes in their investigation of website aesthetics demonstrated that websites that have received a higher level of aesthetic treatment, while displaying the same content, are perceived as being more credible. They

maintain: “*that when the same content is presented using different levels of aesthetic treatment, the content with a higher aesthetic treatment was judged as having higher credibility.*” (Robins & Holmes, 2008). A study to identify the factors of a website that influence credibility was subsequently undertaken by Alsudani and Casey. They found that unity in design, including balance, harmony, contrast, and dominance, were important to positive judgements of credibility (Alsudani & Casey, 2009).

Despite this, there is currently a dearth of research directly investigating the impact of common individual features of a news website’s design, reflected in a news webpage’s aesthetic, on its perceived credibility. The following works demonstrate the piecemeal state of affairs. One of the earliest by Burkell, who previously contributed to one of the aforementioned models, showed that users perceived seals of approval as certifying information quality, which Hilligoss and Rieh claim is closely related to credibility (Burkell, 2004; Hilligoss & Rieh, 2008, p. 1469). Freeman and Spyridakis conducted two studies on the impact of factors including street addresses, links to external websites, and Contact Us facilities. They found that all three positively affect the perception of credibility (K. S. Freeman & Spyridakis, 2004, 2009). Westerwick also demonstrated a link between perceived sponsorship credibility and information assessment (Westerwick, 2013). It should be noted however that when a website’s sponsor had a low credibility reputation, better website design could not improve this. Research by Borah investigating how news frames and hyperlinks influence news credibility and willingness to seek information also demonstrated a positive correlation between the presence of inline or embedded hyperlinks in news articles and participants’ perception of credibility (Borah, 2014). This lends credence to Sundar’s claims relating to ‘information scent’ (Sundar, 2008).

In contrast, Hong, when investigating the impact of message and structural features on the credibility of health related websites found that only message features predicted website credibility (Hong, 2006). However, there are several issues with the design of this study. It did not entail using structural features as within-subject factors, also known as independent variables, to discover whether their inclusion or exclusion influenced perceived credibility. In this study, participants were provided with open access to the WWW and asked to undertake two tasks to search for information on quitting smoking that they would be happy to provide to a family member or friend. Whatever websites each participant decided matched this criterion was then analysed to identify the most common message and structural features. A factor analysis revealed that the most common structural elements were third party endorsements, privacy policy statements, site authorship, site contact information, navigation tools, and membership of the since disbanded *HON network*. Therefore, inclusion in the final data analysis was reliant on the participants and not the researcher. The range of structural features identified in this experiment was limited in scope and does not reflect the features of modern website design.

The two tasks also instructed participants to: “*find information that is of high quality*” for a: “*family member/friend*”. Therefore, participants would have been focused on the textual information contained in each website rather than the overall site or its design. Consequently, it was more likely that participants would have selected websites for inclusion in the factor analysis with strong message quality rather than those with structural features that may impact perceived credibility. Message features are typically analysed by the central or systematic routes, or by the controlled model of human information processing. Structural features are more likely to be analysed by the peripheral route, heuristic strategy, or by the automatic model of human information processing. When the task or information is important, as it was in this instance as it was health related information for a family member or friend, users are also more likely to use the central or systematic routes, or the controlled model of human information processing. When the task is not as important users are more likely to evaluate information by the peripheral route, or heuristic strategy, or by the automatic model of human information processing (Chaiken, 1980; Petty & Cacioppo, 1986; Shiffrin & Schneider, 1977). There is a significant body of research to show that users most often evaluate structural features of the design when they take the peripheral, heuristic route, or by the automatic model of human information processing (Chaiken, 1980; Hilligoss & Rieh, 2008; Petty & Cacioppo, 1986; Shiffrin & Schneider, 1977; Sundar, 2008). A later study by Kim, Park and Bozeman investigating online health information search and evaluation also showed that websites are mostly judged by peripheral cues of source and message credibility (Hyojin Kim et al., 2011). Therefore, the Hong study which focused on message content, which is judged by the central or systematic routes, was unlikely to show that any structural features of a website predict its credibility rating.

The results of a much larger and more in-depth study by Flanagin and Metzger showed that credibility assessments are predominantly formed based upon a website’s attributes, including its features (Flanagin & Metzger, 2007). The 4x2 experiment comprised of fake and real versions of four websites from different genres. All websites showed an identical news story relating to the harmful effects of radiation on pregnant women. The results showed that fake websites with the same design and content but minus the branding had a comparable credibility score to their real counterparts. Thus, they concluded that credibility assessments are primarily the preserve of website attributes such as design features, depth of content, and site complexity, rather than familiarity with the website owner. Although the study did not identify which design features may be responsible for increasing or decreasing perceived credibility, their hypothesis was based on the supposition that visuals, the amount of information on a website, and interactivity might be responsible. Most interestingly the study demonstrated that genre of website had a large impact on perceived credibility. The real and fake versions of news websites and e-commerce websites were rated as significantly more credible than real and fake versions of special interest group websites or personal blogs showing the same story. Explaining their hypothesis, the authors point out that site genera and site attributes may be linked giving the example that news website typically have a more complex layout, interactive capabilities etc.

In one study focusing on health-related websites and in another focusing on online news, Rains and Karmikel, and Chung et al. showed positive relationships between certain structural or technical features and characteristics of a news website's design, reflected in a webpage's aesthetic, and perceptions of credibility. Sundar maintains that such can act as a "*repository of cues*" which may aid credibility evaluation by "*triggering [cognitive] heuristics about the typical nature of the underlying content*" and perceptions of website credibility (Chung et al., 2012; Rains & Karmikel, 2009; Sundar, 2008). In another study relating to how users whose first language is not English assess the credibility of health information online, design aesthetics and functionality were also found to impact judgements of credibility (Young et al., 2016).

Addressing how perceived credibility can be reduced through issues with a website's design and construction, Allport and Pendley demonstrate that overall poor design and surface flaws reduce perceived credibility (Allport & Pendley, 2010). The credibility of online health information has also been shown to be impacted by variations in the presentation of content (W. S. Jung et al., 2017).

Recently, attention has turned to social or crowdsourced methods to determine what features of a news website's design, reflected in a webpage's aesthetic, impact on perceived credibility. Papaioannou et al. developed a recommender system for employing item based collaborative filtering based on specific web page features, while Haung et al. developed a social platform to rate the credibility of webpages (Huang et al., 2013; Papaioannou et al., 2012).

The main conclusions that can be drawn from the research presented above are: 1) There is undoubtedly a link between how information is presented online, especially relating to structural features of a website's design, and its perceived credibility. 2) Not enough research has been undertaken to ascertain which common features, through their inclusion, omission or distortion, can impact upon perceived credibility or any of its measures. 3) Almost no work has been undertaken to determine how overarching design influences can impact upon perceived credibility. 4) Most of the research in the domain does not break its findings down to individual user groups. 5) Due to differences in experimental approach, comparing studies, longitudinally or otherwise, is problematic. 6) The benefits of an empirical framework for credibility assessment would also be of use to assessing the impact of website features and design influences on the perception of bias, or any other dimension and measure of credibility.

2.2.12 Motivating Factors Based on Issues with Current Approaches to Measuring Credibility Online, and any of its Dimensions such as Bias, and the Resulting Requirements for an Open Framework

The following three sections detail the motivations which drove the requirements behind the design and development of the CAFE framework, the repository and classification of measures, and the experiment platform.

2.2.12.1 Framework Requirements

One of the most important requirements of the framework is that researchers of all abilities should find it easy to use, while it supports them in designing and developing experiments to measure credibility or any of its individual dimensions. It is also important that it supports experimental rigour while helping researchers to identify potential issues before the experiment is deployed. A wide range of experiment designs also needs to be supported by the framework. Another motivation behind the development of CAFE is to aid communication between a researcher and other members of the team. This can then act as a record of the experiment plan which can be used in the preregistration of hypotheses and in the post experiment publication of data. It is also important that the framework highlights specific things which researchers can do to improve their experiments such as aiming for a specific statistical power and effect size and including this in the write up. The framework should also highlight the importance of long-term planning, such as the identification of specific publication venues during the planning stages of the experiment.

Based on these motivating issues, the following framework requirements were identified:

- Easy to use framework, which will help researchers plan experiments, formulate their thoughts, consider the implications of their design choices all while helping them to spot potential issues and pitfalls
- Improve experiment rigour and recording
- Help researchers to identify resources and calculate timelines
- The framework must be open and extensible so that it can be adapted to a range of different experiment designs
- Act as a communication aid between researchers, supervisors, and team members to make sure their understanding is in unison
- The framework should act as a record of the planned experiment, capable of being used in the pre-registration of hypothesis and support post experiment research data publication principles such as Findable, Accessible, Interoperable, and Reusable (FAIR)
- Emphasising the importance of reporting statistical power is a requirement of an empirical framework.

2.2.12.2 Repository Requirements

Since the earliest research into the domain, credibility has been measured as a product of its dimensions, the original and foremost being trustworthiness and expertise of the source⁹. Additional source dimensions such as character, competence, sociability, extroversion, composure, safety, qualification and dynamism were put forward in a series of factor analysis studies which found overlapping yet differing dimensions (Berlo et al., 1969; McCroskey, 1966; McCroskey & Jenson, 1975). Metzger et al. provides a synthesis of the dimensions of source, media, and message credibility, their conceptualisations and measurement (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003). More recently, Gaziano and McGrath used similar methods to propose dimensions such as “*is fair or unfair*” or “*is concerned about the public interest or is concerned about making profits*” to conceptualise the credibility of newspapers and television (C. Gaziano & McGrath, 1986). It is only once a dimension is used to measure the concept of credibility that it moves from the theoretical to the empirical, and becomes a measure of credibility, see Figure 2-2.

With the definition of the three main dimensions of credibility, *source*, *medium* and *message*, new and often more appropriate, task specific dimensions, and subsequently measures, have been adopted. They include stalwarts such as fairness, bias and accuracy, but also the less common such as goodwill, sophistication, or a sincere interest in important affairs (Flanagin & Metzger, 2007; Hong, 2006). Consequently, there is no standard set of measures existing in the literature to measure the concept of credibility but rather a situation where researchers must find suitable measures from the literature or invent their own. This creates a situation of the ‘unknown unknown’ where unless each researcher in the domain has done an exhaustive trawl through the literature, they may not be aware of the most appropriate measures for their purpose.

There is also no agreement as to which, how many, or when each measure should be used. Eastin, noting previous research by Gaziano and McGrath suggesting that the choice of measures will influence perception, used three measures to calculate the credibility of online health information; accuracy, believability and factualness (Eastin, 2006). Johnson and Kaye typically used four dimensions, while Sundar has used two, six, and twenty one (Hu & Sundar, 2009; Johnson & Kaye, 1998, 2000, 2002, 2004, 2010; Sundar & Nass, 2001). A single researcher may also vary the amount of measures they use. In one study Chung used fifteen measures of credibility, while in a later study he simply used credibility as a measure of itself (Chung et al., 2010, 2012). The lack of a defined taxonomy or categorization of those used in similar research has made selecting appropriate measures difficult and has thus generated new, unnecessary, and overlapping measures.

⁹ Gunther was one of the few to use an alternative approach, treating credibility as the result of audiences sceptical disposition (Gunther, 1992). Freeman and Spyridakis also point out that Frewer and colleagues used thought listing and interviews in a closely aligned domain, Pennbridge and colleagues collected qualitative responses, and Gaziano and McGrath rated reliability of the media in their study (K. S. Freeman & Spyridakis, 2009; L. J. Frewer et al., 2006; Lynn J. Frewer et al., 1997; C. Gaziano & McGrath, 1986; Pennbridge et al., 1999).

Another issue in the domain is the lack of an abstract or gestalt¹⁰ overview of the range of measures which have been used in peer reviewed research. This will provide inspiration as to other measures that they might not have considered. This will help combat the “unknown unknown” phenomenon, of not knowing about a measure that might be pertinent to their credibility experiment. The repository should also include links to the individual papers so that researchers using it can explore them in more detail. Motivating the development of the repository online is the fact that it can be easily updated and added to. This will help to avoid duplication of existing measures without limiting research

Based on these motivations, the following repository requirements were identified:

- A facility to aid researchers in choosing appropriate measures
- The ability to see which, how many, where and when each measure was used
- Provide a gestalt overview of the measures used in peer reviewed research
- Link to relevant research papers
- The ability to easily add more studies, display additional information about studies, and add new functionality
- Helps to avoid duplication in the identification and testing of additional dimensions without stifling research

2.2.12.3 Platform Requirements

The main motivation behind the development of an open source platform to aid in the design and execution of experiments to measure credibility of any of its dimensions is that it will aid other users without the ability to write code and develop their own experiments from scratch. An opensource platform which can be easily extended and adapted will also greatly reduce the time it takes to deploy experiments for both novice and experienced researchers and for those with and without advanced coding skills. It is also important that the platform supports a broad range of experiment types, including those using Latin squares and cubes and other complex designs.

Another motivation behind the development of this platform are the current issues with many of the scales used to measure credibility or its dimensions online. Typically, researchers in the domain have used Likert or Likert type fixed rating scales, often without knowing the difference. However, even here there is variability, Kiouisis has used 4-point, Johnson and Kaye, and Freeman and Spyridakis prefer 5-point, Flanagan and Metzger, Hong, and Chung use 7-point while Sundar has used 10-point (Chung et al., 2012; Flanagan & Metzger, 2000, 2003, 2007; K. S. Freeman & Spyridakis, 2009; Hong, 2006; Johnson & Kaye, 1998, 2000, 2002, 2004, 2010; Kiouisis, 2001a; Sundar, 1999). Other methods such as dials, semantic

¹⁰ Gestalt is a German word regularly translated to English as pattern or shape, however its true meaning is closer to ‘configuration’. Gestaltism is sometimes referred to as ‘configurationism’. A founding principle of Gestalt psychology is that: “*the whole is something else than the sum of its parts*” (Koffka, 1935). That is, by providing users with an online repository and classification of measures from individual studies in a single table which they can analyse for additional information, provides them with additional insights into the domain.

differential scales, and bipolar adjective rating scales have also been used (S. M. Choi & Rifon, 2002; Ognianova, 1998; Robins et al., 2010; Robins & Holmes, 2008; Z. Wang et al., 2008). Many studies also provide insufficient information as to how these scales have been anchored. Some such as Flanagin and Metzger use mono directional scales anchored with *not at all* and *extremely*, while others such as Hong use bi-directional, *strongly disagree* to *strongly agree* (Flanagin & Metzger, 2000; Hong, 2006). Johnson and Kaye anchored their scales with e.g. *not at all fair* to *very fair* changing the adjective for each measure. Many studies also do not provide detailed information regarding the what instructions each participant received. A high quality research platform needs to be able to provide a range of standard and advanced scales such as VAS.

Another motivation behind the development of an open framework are the increased opportunities for longitudinal comparisons. Early research in the domain was dominated by conflicting studies comparing the credibility of the various mediums, print, radio, television, and the WWW (Flanagin & Metzger, 2000; Johnson & Kaye, 1998; Schweiger, 2000). However, as the studies used differing experimental designs and measures it is difficult to draw overall conclusions. It should be noted that this is not a recommendation that every study should have the same design, method or measures. Differing situations and experimental aims will require different designs, however having a common baseline framework which researchers can adapt to their needs, and increased commonality of measures across different types of studies will remove some ambiguity and allow for better comparisons.

Also motivating the creation of an open framework is the ability to break down a credibility score into its individual dimensions. This will enable longitudinal comparisons of credibility scores and their individual, standardized measures also. While a website medium's overall credibility score may not change significantly, its measures may go through some fluctuation. It will also be possible to break down the score via different participant groups. This will enable researchers to report how various groups of users rate the credibility of a webpage via its dimensions/measures, thus enabling them to ascertain which of the individual dimensions of credibility plays a greater or lesser role among different user groups.

Much of the experimental research conducted in credibility in the past has relied upon student participants (Chung et al., 2012; Flanagin & Metzger, 2000; K. S. Freeman & Spyridakis, 2009; Hong, 2006; Sundar, 1999). Recognizing this, many researchers put significant effort into expanding the sample populations used in experiments through various means, including: written solicitation; magazine subscriptions; and charitable donations; however, these are both time consuming and costly (Flanagin & Metzger, 2003, 2007; Fogg et al., 2001). Consequently, another factor motivating the development of an open empirical framework and platform is the need to be able to take advantage of crowdsourcing, including the recruitment of a much larger, diverse sample, or the ability to target groups outside of Western, Educated, Industrialized, Rich and Democratic (WEIRD) demographics (Henrich et al., 2010a).

Based on these motivations, the following platform requirements were identified:

- Help users without advanced coding skills necessary to develop their own online experiment
- Facilitate quicker deployment of experiments
- Support complex experiment designs using Latin squares and cubes
- The ability to facilitate researchers choice of the most appropriate scale for their study
- Supports longitudinal experiments and/or repeatability in research
- Facilitate the reporting of individual credibility measures
- Support crowdsourced experiments
- Use opensource technologies and make the platform freely available to all

2.2.13 Experimental Frameworks for Detecting Credibility or any of its Dimensions or Measures

An exhaustive search was undertaken of the literature for an empirical framework and/or platform that could be used to undertake the experiments this research necessitated. There were three main phases to this search. 1) Initially it was limited to the domain of bias, as this was the domain in which the experiments would be conducted. However, as this research has demonstrated, there is a complete lack of frameworks in this domain, theoretical or empirical. 2) The comparable domain of trust was the investigated. Although six empirical frameworks for measuring trust were identified, see section 9.8.4 of the appendix, the differences between the domains and the fundamental differences in the nature of the constructs, disqualified them from consideration. 3) Lastly the domain of credibility was examined. Although there is a long history and a broad range of strong empirical research in the domain, no empirical frameworks exist. Consequently, the decision was taken to design and develop an empirical framework and platform, for conducting experiments on the impact of website design on the perception of credibility, or on any of its dimensions and measures such as bias, trust, accuracy, fairness, or depth of coverage.

2.2.14 Conclusions on Credibility

Research on credibility has been conducted since at least the 1940s. Consequently, the domain has been extensively explored which has resulted in a strong foundation of underlying theory. This trend has continued, if not strengthened, since the inception of the Internet and the WWW. This literature review uncovered ten MTFS explaining how users form judgements of credibility online. However, despite an in depth and comprehensive compendium of studies on credibility, no empirical framework has been proposed in the domain. There are several obvious drawbacks to this situation, including easily avoidable deficiencies in the design, development and reporting of studies, difficulties in comparing research, and a lack of common measures for the construct for different dimensions of credibility in different situations. As a result, this research proposes the development of an empirical framework and platform, to be made

open source, which can be used to design and deploy experiments to measure credibility, or any of its dimensions such as bias, online.

2.2.15 Summary on Credibility

This section provided a brief overview of the domain of credibility, of which bias, along with trust (and others), are core dimensions and measures. It highlighted four types of definitions, definitions as a perceptible variable, those based around believability, those focusing on its dimensions, and the philosophical. The two focuses of research in the domain were then delineated, defining and measuring the construct of credibility, and measuring the credibility of source, medium, and message. The meaning of the term ‘dimensions of credibility’, a longstanding issue in the domain, were then explored. For historical context, a classification of credibility assessment techniques was then explored, before the ten MTFS were examined. The underpinning theory of the majority of these, the ELM, HSM, collectively known as Dual-Process Models of Persuasion, and the underpinning C&AHIP, was then explored. They maintain that when users are unengaged, lazy, or the task is unimportant, that they make judgements about the information presented to them by peripheral, heuristic, or automatic information processing routes. This is important, as one of the main suppositions of this research is that the means by which news is presented online, impacts the perception of bias in the news article. These Dual-Process Models of Persuasion underpin the ten MTFS for credibility, and a succession of studies demonstrated the impact of website design on perceptions of credibility. It is the primary hypothesis of this research, that perceived bias in a news article is similarly impact by its visual presentation as bias is a core dimension and measure of credibility, especially when judging news online. It then described related work on the impact of features and characteristics of website design on perceived credibility, of which bias is a core dimension and measure. Issues were then highlighted with current approaches to measuring credibility online, which served to identify the requirements of a planned open empirical framework and platform to aid in the design and execution of experiments. A motivation for the framework, which was based on the requirements, was then provided, which will aid researchers in their investigation of news website features and characteristics on engaged and unengaged users.

2.3 Bias

This section of the State of the Art focuses on bias, which along with trust, expertise, goodwill, dynamism, fairness, accuracy, depth of coverage and others, are core dimensions and measures of credibility. This section will introduce the construct, before highlighting the lack of underlying theory and empirical research. Key definitions of the term from the literature are then provided categorised by their research focus. The impact of bias at the three main stages of the news pipeline, production, dissemination, and consumption, are explored before the impact of bias in the dissemination of news across the three main traditional mediums, print, radio, and television is detailed. The section focuses on bias in the

dissemination of news on the WWW. Finally, the paucity of empirical or theoretical, models, theories, or frameworks which would help in the measurement of perceived bias online is demonstrated.

2.3.1 *Introduction to Bias*

Bias has been a factor in news since time immemorial. Many news consumers consider the news today as being more biased than at any time previously. Yet even a cursory look at the literature will show such claims not to be accurate. Claims and counter claims of bias have been a standard part of the news and politics for generations. Kingsbury and Hart point out that in 1929, C. G. Dittmer asked 150 summer school students at New York University: *“Do you feel that the newspapers are guilty of ‘coloring the news,’ pandering to the interests, and not giving us ‘all the news that’s fit to print?’”*. Out of the 150 students, 119 said yes, 24 said no, and 7 abstained (S. Kingsbury & Hart, 1933). Thus, even as far back as the 1920s nearly 80% considered the news biased. While five Pew research surveys from 1989 to 2012 showed some increase in those perceiving ‘a great deal’ of bias in the news, those perceiving ‘a fair amount’ fell. Overall the perception of bias in the news is fairly consistent (Pew, 2012). In a cutting remark on the reaction of the public and academics to Edith Efron’s seminal study and the resulting best-selling book demonstrating bias in television network coverage of US elections (Efron, 1971), Weaver stated: *“In such an age, the proposition that TV News is biased should be about as controversial as the law of gravity.”* He considered the reaction of many to be a waste of time and resources. He goes on to say: *“Yet the situation is otherwise, to put it mildly. The question is not only asked and debated; it has generated an intense controversy that now includes, in addition to the extremes of left and right, a small army of journalists and network executives, politicians across the spectrum, ethnic and interest groups, presidential commissions, and just about everyone else with political interest or ambition.”* (P. H. Weaver, 1972).

The production, dissemination and consumption stages of the news pipeline are rife with opportunities for its introduction and influence (Park et al., 2009). This pattern has continued in the dissemination of news online, a medium that has continued to grow to the detriment of more traditional formats (Park et al., 2009). Yet despite this, little is known about how the presentation of news online may be impacting perceived bias, or many of the other core dimensions of credibility. With the current focus on the veracity and quality of news, especially online, research is required to understand if the medium and method of presentation may be adding undue influence. Bias, along with trust, expertise, accuracy, fairness, and depth of coverage, are core dimensions and measures of credibility, a foundation stone of quality journalism.

2.3.2 *Bias, The Lack of Underlying Theoretical Research*

Bias is a complex, multifaceted and abstract construct, difficult to describe without context. Claims and counter claims of bias elicit strong emotions and arguments. Often, even discussion of the construct can be difficult as it is frequently brought to the fore during fraught debate on emotive matters. As William so eloquently described the difficulty of discussing the topic: *“Television news bias studies should he*

approached cautiously, since a cool discussion of bias is about as likely as a seminar on women's liberation in a Marine Corps barracks." (Williams, 1975). Many also find it difficult to describe or point to specific examples of bias to which others agree. Whilst many might believe the use of the term bias, whether media bias or political bias, is only recently in vogue due to contemporary political developments, even a surface level examination of the literature will demonstrate that politicians, the media and even the public have made the claim consistently since at least the 1950s. As D'Alessio and Allen wrote in 2000: *"There is probably not an American today who has not heard charges that 'the media' are 'biased.' Many Americans believe that this is true... ..despite the fact that they do not necessarily agree on the nature of that bias."* (D. D'Alessio & Allen, 2000). This statement is even more accurate in today's, politically charged, 'post-truth' era than it was nineteen years ago. Nine in ten Americans now believe that the news is biased (Niven, 2002). As VandeHei & Allen maintain in an article on Politico, about GOP claims of bias in the 2012 US presidential campaign: *"Republicans cry 'bias' so often it feels like a campaign theme. It is, largely because it fires up conservatives and diminishes the punch of legitimate investigative or narrative journalism."* (VandeHei & Allen, 2012). They believe that claims of bias are a deliberate tactic of Republicans to condemn or stricture the media and convince their supporters to dismiss dissenting or unsupportive news.

Despite its importance, there has been a dearth of serious, properly grounded theoretical work on the core construct of bias. As Entman, paraphrasing Niven wrote: *"With all the heat and attention it incites among activists and ordinary citizens, bias is yet to be defined clearly, let alone received much serious empirical attention"* (Entman, 2007; Niven, 2002). This lack of serious empirical and even theoretical work is most evident in the paucity of results when searching the Library of Trinity College Dublin¹¹. As a legal deposit library, it holds more than six million printed works, yet only 64¹² of them have "bias" in the title. This includes several volumes alone from the Bristol Industrial Archaeological Society (BIAS) with the majority of the remaining books on statistical analysis and experiment design. There are no books focusing on explaining or theorizing about the underlying core construct including its constituent dimensions. Compare this to the aligned domain of trust, where there are 1077 recorded works with "trust" in the title, including several dozen books with serious theoretical contributions to the underlying construct including an introduction to bias as a social and cultural issue (Hawley, 2012), a theoretical account of it as a sociological theory (Sztompka, 1999), and a historical overview of the construct (Hosking, 2014).

This disparity is also evident in academic literature. The most cited works relating to bias are all specific to individual biases within a human context. They include contributions on status quo bias in decision making, (Samuelson & Zeckhauser, 1988), negativity bias (Rozin & Royzman, 2001), confirmation bias (Nickerson, 1998) and consensus bias (Ross et al., 1977). The most broad based contribution to the underlying theory is that of Evans who attempted to provide an integrated account of the causes and

¹¹ <https://library.catalogue.tcd.ie/>

¹² Multiple searches conducted between 01/02/2018 and 19/02/2019

consequences of bias in human reasoning (Evans, 1989). These demonstrate the distinct lack of underlying theoretical contributions elucidating on the foundations of the construct or attempts to encapsulate the entire domain. One often overlooked exception to this is the work in social science research of Hammersley and Gomm. They attempted to define bias as one of several potential forms of error including the development of a conceptual framework for identifying types of such bias error (Hammersley & Gomm, 1997). In contrast, there is an abundance of highly cited contributions on the core construct of trust. They include; ‘Trust: The Social Virtues And The Creation Of Prosperity’ (Fukuyama, 1995); ‘Trust As A Social Reality’ (Lewis & Weigert, 1985); ‘The Meanings of Trust’ (McKnight & Chervany, 1996); ‘Can We Trust Trust?’ (Gambetta, 2000); ‘This Thing Called Trust’ (Stoneman, 2008) and; ‘Not So Different After All: A Cross Discipline View Of Trust’ (Rousseau et al., 1998).

This distinct lack of underlying theoretical research on bias should be kept in mind when attempting to understand the construct. It brings to mind, the oft recited but miss quoted phrase: *‘I know it [pornography] when I see it, and this is not it’*¹³, which has entered the English lexicon to categorise an observable, subjective fact or event that lacks clearly defined parameters. Noting the difficulties in designing studies into bias lead Sachsman to conclude: *“bias has generally been defined as what experts (coders or judges) think is bias”* (emphasis on ‘*what experts*’ and ‘*think is bias*’ in original text) (Sachsman, 1970). He also noted that there was no agreement between experts and the general public, on what constitutes bias and recommended this as a future line of enquiry. This literature review has yet to discover such. To further complicate investigations into the subject, most people also believe that they can recognise bias and that they are immune to its effects. Ironically, this is a well-known cognitive bias, called Illusory Superiority bias, a common effect of which is that the majority of drivers believe that they are above average (McCormick et al., 1986; Svenson, 1981). It can also be known as Bias Blind Spot in which the impact of bias in others is recognised, yet its impact or influence in one’s own judgement is not recognised (Pronin et al., 2002). Even experts in the domain find it nigh on impossible to list and categorise all forms of bias, never mind explain them. This is likely one of the factors which has contributed to the paralysis in research to tackle the underlying psychosomatic influences or to encapsulate the domain. However, even the existent literature is sparse on reasons for the lack of underlying theory. As Entman proclaims in the opening paragraph of his highly cited work on the subject: *“The organizing concept is bias, that curiously undertheorized staple of public discourse about the media.”* (Entman, 2007).

¹³ The original or accurate quote is taken from Mr Justice Stewart of the US Supreme Court concurring judgement, when it reversed the decision of the Supreme court of Ohio in a case relating to obscenity and censorship. The full quote reads: “I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description [“hard-core pornography”], and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that.” (Jacobellis v. Ohio, 378 U.S. 184 (1964), 1964).

2.3.3 Definitions of Bias

Definitions of bias fall into two categories, cognitive and domain specific. Cognitive biases are psychosomatic in nature and their definitions are designed to diagnose biases within an individual. These biases can be sub divided into categories such as decision making, belief, behavioural, social, and memory error biases. However, definitions of cognitive bias are out of scope for this research. Domain specific bias definitions can be general in nature, such as definitions of political bias or media bias, or more precise, such as definitions of bias in the favourability of photographs, framing, or agenda setting biases. Both cognitive and domain specific definitions of bias have typically been put forward to diagnose and subsequently to measure their impact. However, this process is fraught with difficulty. As Park et al. stated, echoing many others in the domain: “*As described, it is intrinsically difficult to objectively define what bias is and so is to measure or correct it*” (Park et al., 2009).

2.3.3.1 Difficulties in Measuring Bias

There are seven main reasons for this. 1) The lack of an agreed formal underlying theory of bias; a baseline definition of the term; formal, tested human-centric theories as to how judgements of bias are formed; and the lack of serious centrally organised, and abutted, repository of empirical academic research. By centrally organised we mean organised and conjoined studies on different facets of the construct. Historically, the majority of research has been conducted in isolation. The resulting patchwork of studies means that a large body of evidence on any particular facet of the construct is lacking, thus making it difficult to formulate more general underlying theories. 2) Overlapping biases are also a major issue in the domain. Often the differences between biases such as; source; selection; bad news; framing; narrative; and partisan bias, are subjective and have large areas of overlap. In some instances, they are grouped together as different facets of ‘reporting bias’, or even under the more general terms, ‘news bias’, or ‘media bias’. 3) To confuse matters further, bias must also be put in context. In the lifecycle of a typical news story, every actor in the process may suffer from one or more of the same biases. However, they may manifest and impact news differently, e.g. confirmation, selection, agenda setting, and bandwagon biases may impact the subject, journalist, editor, news agency, and consumer. It is therefore necessary to contextualise individual biases at the different stages, e.g. partisan bias in the journalist, confirmation bias in the consumer etc. 4) Bias is subjective. This problem has been well documented, most notably by Groeling and earlier by Groeling and Kernell (Groeling, 2013a; Groeling & Kernell, 1998). They maintain that different perceptions of news might arise because of differences in article content and previous experience. They also maintain that cognitive biases within the consumer will impact the perception of bias. 5) Groeling has also defined the problem of the ‘Unobserved Population’ in studies of media bias. He maintains that any research needs to contend with the fact that published news articles are only a small selection of the news: “*The issue of selection bias presents this research with a serious conundrum. How can it assess the representativeness of the sample when the population is comprised mostly of stories that were never reported and thereby elude observation?*” (Groeling, 2013a). 6) Many biases contribute to, or are factors in, other biases.

Selection and/or source bias in journalists are often two contributing factors of partisan bias. There is also an argument that selection, partisan, agenda setting, or framing biases are the result of ideological biases which can be traced back to underlying cognitive biases in the reporter or editor. 7) Lastly, there is a lack of an agreed compendium or classification to define, delineate, diagnose and record different biases, their manifestation, affected domains, impact, or reception.

These issues are not limited to news biases. There is also no agreed classification of cognitive biases.

While D'Alessio and Allen, Entman, and Park et al. are all regularly cited as describing the issues of lack of agreement of what exactly bias entails, the lack of agreed definitions, and the lack of serious underlying theoretical research, they had the benefit of hindsight and a plethora of research demonstrating this point (D. D'Alessio & Allen, 2000; Entman, 2007; Park et al., 2009). It was Edith Efron, who was one of the earliest to highlight these problems clearly and to forge ahead with a solution. In a manner that belied the complexity of the task before her, she modestly stated: *"There is only one way to solve it—and that is to arrive at a clear and demonstrable definition of political bias, to define a simple analytical method for ascertaining the presence or absence of such bias, to apply this method to the network product, and to arrive at a documented answer."* (Efron, 1971).

However, it was Lowry's subsequent analysis of the problem that clearly defines the difficulty of measurement while putting forward a solution. He points out that bias may only be measured when two conditions are met: *"The first condition is that one must be willing to accept a relative definition of bias and must recognize that it is impossible to measure bias in human communication in any 'absolute' sense. Therefore, bias in news reports and any other kind of human communication must always be subjectively defined. However, once it is so defined, it may be objectively measured within the context of the definition"* (D. T. Lowry, 1973).

2.3.3.2 Bias Criteria: Volitional, Influential, Threatening, Choice, Sustained, Systematic and Effective
Several researchers have attempted to define bias by applying identifiable labels or by questioning whether it meets certain criteria. Many of these criteria subsequently then ended up as the basis for definitions of bias used to measure the construct. D'Alessio and Allen highlight the work of Williams, who in his comparison and review of four early studies on bias in television news, placed four criteria on whether bias exists. He maintains that bias must be volitional or wilful, putatively influential, threatening to conventional values, and it must be sustained long enough to be effective (D. D'Alessio & Allen, 2000; Williams, 1975). More recently, researchers do not place as much weight on whether bias is volitional or wilful, recognising that it can be unconscious or passive. One definition of bias in the aligned domain of social science research even specifically excludes motivation, preferring to define it in terms of systematic and culpable error, thus enabling them to: *"distinguish between motivated and unmotivated bias"*, see Figure 9-1 in section 9.3 of

the appendix (Hammersley & Gomm, 1997). Most other researchers choose to ignore motivation as a factor, such as Brandenburg's conceptualisation of the construct for a study of bias in the coverage of the 2002 Irish general election. He states: "*The notion of bias implies a deviation from the informative media function, which may result in a distorting effect on political attitudes and outcomes.*" (Brandenburg, 2005). Like many of the definitions that are detailed in the following sections, this notion of bias places no weight on whether it was wilful or volitional. Instead, more recent definitions of bias tend to focus on other criteria such as whether it is systematic or effective.

2.3.3.3 *Definitions of Cognitive Bias*

Cognitive biases are perhaps the single largest research area relating to bias and as such are almost completely within the purview of psychology and/or psychiatry literature. However, there is an argument that bias in the news, real or perceived, conscious or unconscious, is the result of either cognitive biases in the communicator or in the receiver. There is also an argument that as cognitive biases play a not inconsequential role in this research, an accepted definition from the literature should be forthcoming. As such, a general definition of cognitive bias accepted here is that by Williams, put forth as part of his review of four early books on the subject of bias in television networks: "*By whatever name, bias implies a crucially threatening deviation from norms central to cognitive balance.*" (Williams, 1975).

2.3.3.4 *Non-Definitions of Bias*

In 1981 Kline wrote: "*Still a comprehensive definition of bias remains an illusive goal. Bias research which originally sought merely to illustrate how particular interest groups influenced the partisan political process through the news has revealed a complex variety of institutional, organizational, technological and psychological factors which give shape to journalistic decision making*" (Kline, 1981). The use of the term 'illusive' rather than 'elusive' is noteworthy. Illusive means not based on reality or something that does not really exist, while elusive means difficult to catch or define. This statement has proven more accurate with each study conducted in the nearly four decades since it was written. Instead of the domain solidifying around a central definition, it has increasingly fragmented. This can be seen in the various definitions presented in the following sub-sections. This issue has also led some researchers to refrain from measuring the construct at all. Deacon, Golding, & Billig explicitly abstain from measuring bias. They state, rather erroneously: "*bias in the news is notoriously difficult to define and, for this reason, analysts avoid using the term*" (Deacon et al., 2001).

Alternatively, some researchers believe that a suitable definition of bias is impossible and substituted other easier to define measures in their studies. This is evident in the work of Druckman and Parkin who instead investigated 'editorial slant' as a substitute. They state: "*One of the prime lessons of research on media slant or bias concerns the futility of searching for an 'objective' standard by which to assess bias.*" (Druckman & Parkin, 2005). In another example, Forward also agrees that a definition of bias is impossible

and instead investigated favour and disfavour. He claims: *“a workable definition of ‘bias’ is impossible to obtain for newspaper treatment of the government of the day, with omissions being the main stumbling block. This present study is therefore not about bias, but merely about manifest favour and disfavour.”* (Forward, 1977).

Several researchers in the domain connect the terms ‘loaded’ or ‘loading’ with bias (Stoodley, 1960). Sachsman uses this to create a new measure of bias. Noting the fact that there is no agreement between what experts believe is biased and what the public believe is biased, he maintains that: *“bias has generally been defined as what experts (coders or judges) think is bias”* (emphasis on ‘what experts’ and ‘think is bias’ in original text) (Sachsman, 1970). In an almost comical avoidance of defining either term, Sachsman maintains: *“‘Loading’ is defined as what members of the general population see as loading. Thus loading=slanting=biasing.”*

The reticence to define bias is also true in some of the early, formative, and seminal works in the domain including four early books on the subject (Efron, 1971; Epstein, 1973; Frank, 1973; Lefever, 1974). As Williams points out in his review of these studies into television news bias that none offer a real explicit lexically acceptable and tested definition of the term (Williams, 1975). It should be noted that Efron used the FCCs definition.

2.3.3.5 *Domain Specific Definitions of Bias*

A broad range of definitions of bias can be found in the literature. These range from semi-broad definitions of media bias, to specific definitions of photograph or headline bias. To better understand and appreciate this range, this section presents some of the most cited and relevant definitions divided into nine sub-categories. It does not include definitions from several other large areas of literature such as bias in statistics, sample selection bias, or engineering bias which are unrelated. Of course, there are large areas of overlap between the below categories, and much debate could be had as to which definitions belongs in which, yet this breakdown is offered as the first of its type in the domain.

In each of the categories the definitions are presented in the order they were made public. It should be noted that many of the studies, especially the older formative works in the domain, do not actually use or put forward a specific lexical definition of the term, but rather offer what could be best described as a description or conceptualisation. This is due to the aforementioned difficulty in crafting a suitable lexical definition and possibly due to a reluctance to put forward a poor crafting of words or to being tied to it. In the interest of thoroughness, many of the most important of these descriptions or conceptualisations are also presented. The following categorisation the single largest collection of definitions, descriptions and conceptualisations found in the literature.

For brevity, these categorisations were moved to the appendix. They include: Definitions of Media Bias (appendix 9.2.1.1.1), Definitions of Political Bias (appendix 9.2.1.1.2), Definitions of News Bias (appendix 9.2.1.1.3), Definitions of Political Bias in News (appendix 9.2.1.1.4), Descriptions of Content or Photograph Bias (appendix 9.2.1.1.5), Definitions by Differentials in Coverage (appendix 9.2.1.1.6), Categorisation of Definitions of Bias by Related Terms (appendix 9.2.1.1.7), Algorithmic Definitions of Bias (appendix 9.2.1.1.8), and Definitions of Bias in Social Science (appendix 9.2.1.1.9).

There are four overarching points apparent from this categorisation of definitions:

- There is no, and may never be, an agreed underlying definition. This is likely the result of the lack of formal underlying theory and research to underpin the domain. This has resulted in researchers crafting an ever-expanding array of definitions specific to their research needs.
- Each definition is dependent on the domain and is subjective. As per Lowry's analysis of the problem, which is described in section 2.3.3.1, each definition is subjective and relative to its domain, and bias can only be measured within the context of that definition (D. T. Lowry, 1973).
- There is a lot of dissimilarity even within narrow domains. One of the most surprising aspects of this categorisation is the lack of similarity in definitions of bias within small sub domains such as Definitions of News Bias (appendix 9.2.1.1.3), Definitions of Political Bias in News (appendix 9.2.1.1.4) or any of the others in the appendix.
- There are no definitions that focus on the impact of the presentation. Even within the related domain of Visual Content or Photograph Bias (appendix 9.2.1.1.5), the majority of studies do not offer a specific lexical definition, but rather offer a broad description using terms like 'warmth'.

2.3.4 Bias in the Three Main Stages of the News Cycle: Production, Dissemination, and Consumption

Bias is a consistent factor in the production, dissemination and consumption of news. While this fact appeared surprising to many in the early years, some scholars found the reaction to this knowledge as somewhat preposterous. As Weaver wrote in response to Efron's early seminal study highlighting bias in network coverage: "*In such an age, the proposition that TV News is biased should be about as controversial as the law of gravity.*" (Efron, 1971; P. H. Weaver, 1972). Much of the early work during the 1970s went about proving and re-proving that much of the news was in fact biased, with the US presidential election and the typical two party, two candidate nature of their political system serving as a handy and regular test environment. During the following decades, much of the work focused on identifying specific incidents of how bias creeps into the news pipeline or the impact it has on the consumer. More recently there is also increasing awareness of the need to protect the consumer, and a desire to develop means to combat bias, most notably the attempts from Park et al. (Park et al., 2012, 2009; Park, Ko, Kim, et al., 2011; Park, Ko, Liu, et al., 2011). Despite this, there is no doubt that there are still many sub-stages in the news pipeline which have received little or no serious empirical study. While this is slowly being

ameliorated, there is now a sense that bias has become an increasingly dirty word, an invective used by those who do not agree with or trust the news, or by those who simply have an axe to grind. Consequently, claims and counter claims of bias are now used to dismiss or diminish the value of news. Because of this, even appropriate or accurate use of the term can result in accusations of fearmongering from those who either disagree with a certain stance or position or who have dismissed it as a serious issue due to hearing the cry of wolf all too often. Despite this, it is still necessary for the research to continue.

The following sections are a first attempt to categorise as many of the different biases which impact on news as possible. This is a complex task, yet the fundamentals of the situation are that imbalance in any stage or aspect of news coverage can be considered a bias.

The following are the main issues with classifying bias. 1). In some instances, the type of bias is specific with established nomenclature and examples in the literature, such as selection bias and partisan bias. In other instances, they are much less specific or well known, such as class bias (McChesney, 2003) or description bias (Earl et al., 2004). Some authors also use the same terms to describe two different biases or even more confusingly, describe the same bias with two different terms. 2). They are huge areas of overlap. 3). Domain affiliation or attribution can add another level of complexity. This can range from the general term media bias, to the narrower term news bias, or to specific event terminology such as election coverage bias. The same bias can be prevalent in each instance, e.g. partisan media, partisan news, or partisan election coverage bias. 4). Different biases are not limited to different actors, nor are they limited to the same actors. News agencies, editors, journalists and consumers can all be guilty of selection bias, yet only news agencies, editors and journalists can be guilty of agenda setting bias. 5). Individual biases enable and contribute to other biases. One of the most common means of framing bias is to select certain facts that support or oppose a position. Thus, selection bias contributes to framing bias. 6). Domain, Agents, Events and/or Supertypes. To confuse matters still further, bias is often attributed to or referred to in terms of the agents, events and/or supertypes. Examples of domain bias include conservative radio or liberal talk shows in the US. Agent biases include journalistic bias or source bias. Event bias examples include election bias, debate bias, or sporting coverage bias. Lastly supertypes have also perforated throughout the domain, examples include reporting bias, journalistic bias, or newspaper bias, all of which are overarching terms for a collection of one or more biases in the agent or event covered. Table 2-2 represents a small proportion of domain, agent, event, and supertypes terminologies. These issues and others have recently been explored as part of an initial classification of biases impacting the production, dissemination and consumption of news (Spillane & Wade, 2020).

Table 2-2 Examples of Domain, Agent, Event, and Supertype categorisations of bias.

Domain	Agent	Event	Supertype
News	News Agency	Election	Reporting
Media	Television News Network	Debate	Journalistic
Newspaper (print media)	Newspaper	Sporting Occasion	Editorial
Radio	Radio Station	Referendum	Newspaper
Television	News Website	Court Case	Conservative
Online News	Journalist	Social Movement	Liberal
Government	Editor	Political Event	Television
Corporate	Owner		Visual
Religious	Consumer		Presentation

2.3.4.1 *Bias in the Production of News*

The production of news is a complex process with a constant race against deadlines and competitors. Stories and articles can take days, weeks, months, or even years to produce. Despite this, television, newspapers, and radio all work to tight deadlines and fine margins. News becomes less valuable as a commodity the older it is and as competitors publish in advance. While it might be argued that such finite deadlines and margins do not exist for online news, the increasingly multimodal nature of news websites means that they are now often integrated with live news bulletin from a traditional media source or a live stream of an event such as a press briefing, or election results. No matter the medium, news agencies are always focused on beating their competitors to the scoop. If credibility and its individual dimensions such as accuracy and trustworthiness are the most important aspects of a news organisations reputation, speed is most definitely the second most important. Many would even argue the reverse. News websites push to be first with breaking news has begun to have a detrimental effect on their credibility. Kalev Leetaru, a senior fellow at Georgetown University writing for RealClearPolitics maintains that: *“the media’s prioritization of speed over accuracy is one of the reasons it has lost so much credibility”* (Leetaru, 2019) Only a small number of news organisations, predominately news magazines with a monthly or bi-monthly deadline, or news organisations such as Euro News,¹⁴ have the luxury of not being regularly or hard pressed to such deadlines.

News organisations live and die by the speed of their access to unfolding events. While news websites do not have the hard deadlines of radio, television, and especially print, they are expected to publish an almost continuous stream of new content. This is most prevalent in their news apps, which like their news websites, are just another medium for content from their news publishing platforms. With the advent of push notifications, news websites are now increasingly in a race with each other to provide almost instantaneous updates about an unfolding event. Often the pressure to be the first to produce and publish such news means that articles are pushed live while the events are still unfolding, with a notification to such effect posted at the bottom of the page. In many instances, events are now reported in real time as part of a live event page, continuously updated as events unfold. The real-time nature of modern online news means that many of

¹⁴ www.euronews.com A European wide news agency which aggregates content from its local shareholder news outlets around Europe and concentrates on factual analysis and accuracy rather than the speed or currency of its content.

the checks and balances and reflective time that was available in more traditional mediums is no more. This includes extensive fact checking, editing, review and other preventative measures to reduce bias.

Despite its importance as a topic, bias in the production of news is out of scope for this thesis. However, a full and detailed treatment of the topic can be seen in appendix 9.3.

2.3.4.2 Bias in the Dissemination of News

Bias in the dissemination of news has been studied since at least the 1960s. Dissemination is here defined as the means of delivery of the intended message. In newspapers, it includes; all textual and visual content. It also includes amount of coverage, headlines, and layout and spread of articles. In radio, it includes; the spoken word, choice of words, level of respect, and rhetorical formats such as use of sarcasm. It also includes; the tone of voice, and other vocal characteristics such as rate, intonation, pitch, prosody, cadence, and accent. It also includes the amount of coverage, such as the duration, prominence, and repetition of segments. Bias in the dissemination of news on television includes; the seriousness of attention, the spoken and written word, the tone of voice, prosody, choice of words and interview style. Like radio, it also includes the amount of coverage.

Although important in their own right, for brevity this section was moved to appendix 9.4 It includes detailed breakdowns of relevant research on of bias in the dissemination of news on the radio (appendix 9.4.1), bias in the dissemination of news in print newspapers appendix 9.4.2), and bias in the dissemination of news on television (appendix 9.4.3).

2.3.4.3 Cognitive Bias in the Consumption of News

There is no doubt that there are many cognitive biases present within the consumer and active during the consumption of news. Their impact can range from mild to severe depending on the individual or subject matter. While research from a predominately, psychology, psychiatry, social science prospective, or from within the fields of communication studies. However, they are out of scope for this thesis.

2.3.5 Bias in the Dissemination of News in Traditional Mediums: Print, Radio, and Television

The following sections provide an overview of work on bias in the dissemination of news in traditional mediums. The main work has been moved appendix 9.5, 9.6, and 9.7. To read an extremely in depth review of research into visual communication in political communication, including a detailed review of related work on bias, Scill's "The Visual Image and The Political Image" not only provides an alternative overview, but also discusses the main functions of visuals in politics (Schill, 2012). Gunter's book "Measuring Bias on Television" is also a significant reference point for studies in the domain (Gunter, 1997).

2.3.5.1 Studies on Bias in Print Newspapers and News Magazines

There have been more studies into bias in the dissemination of news in print media than any other medium. A classification of the different studies is provided in Table 9-7 in the appendix. An in-depth review of five of these areas of study is also provided in the appendix. These include a significant number of studies on; coverage bias (appendix 9.5.1), headline bias (appendix 9.5.2), photograph bias (appendix 9.5.3), frontrunner bias (appendix 9.5.4), and the supertype visual bias (appendix 9.5.5).

2.3.5.2 Studies on Bias in Radio News

Compared to the other traditional mediums, there have been very few studies on bias in the production or dissemination of news on the radio. A detailed description of the reasons for this is provided in section 9.6.1 of the appendix. As a result, the review was only able to identify a smattering of studies into bias in the dissemination of news on the radio, these include; studies identifying partisan and class bias on Australian radio (appendix 9.6.2), studies identifying partisan or pro-government bias in South African broadcasting (appendix 9.6.3), new methods and techniques for identifying bias in speech in radio broadcasters (9.6.4), favourability and coverage bias in radio (appendix 9.6.5), a lack of research into the influence of advertising on radio bias (appendix 9.6.6), and studies identifying incumbency bias in radio (appendix 9.6.7).

2.3.5.3 Studies on Bias in Television News

There has been a significant amount of research into bias in television over the last five decades. Most of this work has been undertaken on organisational level bias, such as research into the ideological partisanship, particularly political partisanship, of different television networks, studies on gatekeeping among different networks and issues surrounding race. However, there has also been some significant contributions on bias in the dissemination of news in television. The following paragraphs highlight some relevant research in this area, a wider overview is provided in Table 9-8 in the appendix. It must be noted that there are large overlaps in the types of dissemination bias present in television news. This is mostly due to the multi-modal nature of the medium where for example, news headlines may be listened to in audio format, watched in visual format, or read from the screen, or some combination of the three.

A detailed overview of studies into four categories of bias in the dissemination of news on television includes; studies into presentation bias in television news (appendix 9.7.1), studies into bias in facial expressions of news anchors (appendix 9.7.2), studies into soundbite and imagebite bias in nightly news shows (appendix 9.7.3), and research into the supertype visual bias in television News (appendix 9.7.4).

2.3.6 Bias in the Dissemination of News on the WWW

The previous sections highlighted research on bias in the dissemination of news in traditional mediums, print, radio, and television. The following sections highlight a selection of the main contributions on research on bias in the dissemination of news on the WWW to demonstrate the breadth of the domain.

Table 2-3 displays the categorisations of bias which may impact the dissemination of news on the WWW, but the vast majority of which have yet to receive any serious empirical attention from bias researchers. It should be noted that many of these categories of bias have received particular attention from other research areas, some of which are related. Online news effectively blurs the line between print and television, in that it combines, text, images, video, and to a lesser extent, interactive features. It also often contains the same news articles or news videos from the news agencies print publications or television stations. Consequently, most of the categories in Table 2-3 below are taken from Table 9-7 and Table 9-8 in the appendix, yet despite an exhaustive search, very little, or even no peer reviewed academic research could be found for most of them that was undertaken from a bias prospective. N/A indicates where no research was identified. The purpose of Table 2-3 is to show the huge disparity in the volume and depth of research between the domains. It should be noted that the research presented in Table 9-7 and Table 9-8 represents only a small proportions of the most relevant studies for each categorisation, while the studies identified in Table 2-3 represent all the research that could be found for each categorisation.

Table 2-3 Categorisation of biases that have been investigated in the dissemination of news on the WWW, their manifestation and effect, including references to relevant papers where found.

Type of WWW Dissemination Bias	Manifestation / Effect	Paper
Position, Placement, Prominence	Placement of stories in a newspaper such as front page or even 'above the fold' or buried deep within the paper. Position or placement bias can also manifest in the placement of facts within a story	N/A
Layout, Spread	Placement and groping of articles, photos, headlines to create an initial impression	N/A
Coverage, Space	The amount of coverage of a topic, issue, political party or individual receives. Closely related to position / placement / prominence	(Chakraborty et al., 2015)
Visual	The visual image presented by a news anchor or news agency, of an individual, candidate, party, or organisation to communicate a favourable or unfavourable representation or frame a story	(Goodnow, 2010; Kwak & An, 2016; Narwal et al., 2017; C. Robinson, 2016)
Photograph	Favourable or unfavourable converge of an individual or event, or the attempt to convey a particular impression, by presenting an individual or event in a certain light or from a certain viewpoint.	(Hehman et al., 2012; Towner, 2017)
Graphics	Biased reporting of information by emphasising or highlighting certain data to support an argument	N/A
Description	Word choice, use of titles, tone, focus, respect, endorsement or condemnation which may contribute to erroneous reporting of events	N/A
Labelling	Use of titles, nicknames or other terms of disrespect. Highlighting certain individuals background, religious or political beliefs	N/A

Symbols, Symbolism	This can include the use of iconography or symbols to represent an organisation in a certain manner the continuous use of certain visual representations to depict a certain group, race, culture or religion	N/A
Presentation (media)	An overarching term to describe news media bias. Often involves the positive or negative representations of an individual, situation, or organisation by a number of visual, linguistic, or other means to create a reality different from the truth	(Bar-Ilan et al., 2009) (Groeling, 2013a)
User Interaction or Comments	One-sided display, moderation, or voting patterns among comments in a news forum	(Houston et al., 2011)
Soundbite, Imagebite	The choosing of sound or image clips, often as an advertisement to a larger story, which may frame or set the agenda for a story or make an individual or cause appear in a favourable or unfavourable manner	N/A
News Article Format	The use of different story formats, such as balanced view, or inference story, or discussion formats such as lecture, interview, or discussion, to emphasise different aspects of a news story or to frame it in a certain manner	N/A
Program Format*	Differences in the formats of new news shows such as cable news, political satire, and late-night news talk shows on the news they produce	N/A
Factual Presentation in the Message	The presentation of certain 'facts' about a news story. What constitutes a 'fact' may be subjectively interpreted by the news anchor. They may also provide unbalanced coverage of such facts	N/A
Editorial or Journalistic Opinion in the Message	The knowing or unknowing insertion or addition of journalists or news editors' personal opinions to slant the news	N/A
Objectivity in Message	Was the message or content objective in terms of time, neutral voice, objective language, evidence based, and given equal time in the news pipeline	N/A
Language, Word Choice, Level of Respect, Rhetorical Formats such as: Sarcasm, Hyperbole, Sneer, Sardonicism	The use of biased language to identify or differentiate against a particular group or cast them in a disparaging light	N/A
Facial Expressions*	Perceived bias in the facial expressions of news anchors which along with other non-verbal cues can communicate scepticism, disdain, disgust, sarcasm, etc.	N/A
Personal Presentation and Clothing	The personal appearance of the news anchor which has been shown to make a message to be perceived more credible and less biased. It also includes the personal appearance of the subject of the news which can range from prison uniforms to three-piece suits	N/A
Physical Characteristics and Attributes	The impact of physical characteristics of the source, typically a news anchor, on the perception of bias. Typically measured as part of credibility	N/A
Audience Response Information	Unbalanced or slanted accompanying text and graphics which may be overlaid on the news such as viewer polls	N/A
Vocal Characteristics* such as: Speech Rate, Intonation, Tone, Pitch, Prosody, Cadence, Accent	The use of paralinguistic elements of speech to influence the content of the message being conveyed to the consumer	N/A
Headline, Lead, Lead Story, Bulletin,	Use of words, tone, focus, respect, endorsement or condemnation to frame the readers thinking whether they read the article or not	(Weatherly et al., 2007)
Favourableness or Degree of Polarity +/-	The degree to which statements made in a television news broadcast are favourable / unfavourable to one candidate, party, ideology or position	N/A

of the Content or Message on a Scale		
Nonverbal Communication*, Gesturing	Communicating nonverbally in a favourable or unfavourable manner towards a person or situation	N/A
Chyrons	Refers to the captions or graphics that occupy the 'lower third' area of a television screen during a news broadcast.	N/A
Accompanying Graphics	The presentation of graphical information to support one side in an argument	N/A

* Included in this list as many news websites, e.g. Reuters, Al Jazeera now have online news shows

Bias in the dissemination of news is traditionally thought of in terms of textual content or the spoken word. Consequently, many studies in the domain undertake a content or rhetoric analysis. These include studies to measure the positive or negative statements, comparison studies of content, or an analysis for specific incidents of terms which indicate allegiance to different ideologies (Groseclose, 2011; Latham, 2013).

However, the vast majority of studies in the domain do not undertake a content analysis studies and instead focus on other factors surrounding the content which are part of or facilitate its dissemination. As Pride and Wamsley stated nearly five decades ago in their study comparing audio and audio visuals in news reports for bias during the incursion of Laos: “*Clearly one cannot determine bias by content analysis alone.*” (emphasis on ‘bias’ in original text) (Pride & Wamsley, 1972). The research undertaken as part of this thesis has also focused on the means of delivery of the content, in this case the visual presentation of news on the WWW. They go on to say: “*Content analysis does not permit one to evaluate directly the appropriateness of the content against reality.*” That is to say, in a flat analysis of the content, elements such as the tone of voice, which can indicate scorn, sarcasm, deference or respect towards experts or guests, are difficult to detect. A content analysis done in isolation will also not take into account its appropriateness against the reality of the situation. Judging the content of a single news report in isolation is not indicative of bias in the news agency or in the media in general. Hence Pride & Wamsley’s decision to analyse 30 days of news reports by CBS and NBC about the US incursion of Laos.

2.3.6.1.1 *Headline Bias in News on the WWW*

Just as they were in print media, headlines have been investigated for bias in news on the WWW. In their study of political bias in the headlines used by CNN and Fox News during the final two months of the 2004 presidential campaign, Weatherly et al. found that CNNs headlines were significantly more liberal than those taken from Fox, and somewhat surprisingly, Fox news headlines were also rated as slightly to the left of neutral (Weatherly et al., 2007). It must be noted though, that other than this example, the only other examples of recent studies into bias in headlines in online news are all master’s thesis or undergraduate research (McCluskey, 2005; Sexton, 2014; Simpson & Curtis, 2011). In a somewhat one-sided analysis of western press contribution and involvement in the War in Iraq, Muscaati maintains that: “*The use of headlines can also serve to instil bias. For many fast readers, headlines become their only source of news*

on many issues. Even for those reading an entire story, headline bias can form a contextual barrier within which events in the story are framed.” (Muscati, 2002).

2.3.6.1.2 Coverage Bias in News on the WWW

While coverage bias has been extensively studied in traditional media, especially print media, there have been relatively few empirical studies online. One notable exception is that of Chakaborty et al., who looked at whether trending news stories on news and social media websites could create coverage bias. The premise for the study is their assertion that: *“Due to the round-the-clock (24/7) nature of online news and the need to keep their audience coming back to their site, online news sites today are emphasizing recent news stories over relevant or important news stories.”* (footnote in original quote) (Chakraborty et al., 2015). Their study found that churn in popular content can create temporal coverage biases based on the time of day the user accesses the websites. Coverage bias has also previously been studied on social media websites alone by Chakraborty et al., where a growing proportion of users are receiving their news (Chakraborty et al., 2016), in comparisons of Russian state and independent news agencies (Watanabe, 2017b), and comparisons between cable news and online news websites (González-Bailón et al., 2014).

2.3.6.1.3 Photograph Bias in News on the WWW

In a sign of the increasing importance of the WWW as a medium of news dissemination, research has begun to focus on bias in the photographs which accompany online news articles. Hehman et al. undertook a content analysis of photographs depicting American presidents Bush and Obama for warmth and competence (Hehman et al., 2012). The study showed that media outlets that shared political ideologies with the candidates used imagery which showed the candidate with more warmth and competence than candidates they did not share political ideologies with. Recently, the role of bias in visual content in social media in US elections has been the focus of increased scrutiny. Towner found that visual campaign content such as images and infographics has more of an effect than textual campaign content on social media (Towner, 2017).

2.3.6.1.4 Presentation Bias in News on the WWW

Presentation bias is a supertype, see Table 2-2, first coined by Lin et al. in a study of bias in television news reporting (W.-Y. Lin et al., 2011). Groeling also used presentation bias as a catch all term, claiming the supertype: *“is the focus of the vast majority of the media bias literature”*. It should be noted that this literature review did not find a single empirical study specifically relating to “presentation bias”. Nevertheless he states: *“I define partisan presentation bias as composing news stories in a manner that presents a significantly distorted view of reality, which systematically and disproportionately favors one party over the other.”* (Groeling, 2013a). The impact of presentation bias is regularly noted in studies of search results, many of which focus on news search tasks. Presentation bias is sometimes used instead of

the more common term, position bias. Bar-Ilan demonstrated that presentation bias in the placement of results, particularly their order, was an important factor in user choice. (Bar-Ilan et al., 2009).

2.3.6.1.5 Visual Bias in News on the WWW

Compared to print or television, visual bias is receiving increased attention online, but in a different form than is has traditionally been investigated. This is due to the format and amount of content, which often makes it more readily available and suited to computational methods to undertake laborious and time-consuming tasks such as the data collection and coding of imagery. Kwak and An used deep learning to analyse millions of news photos for visual bias relating to sentiment of photos, whether they are aligned with the sentiment of the accompanying news article, and to investigate how different candidates are treated. Among their many findings in this complex study, was that pro Clinton media tended to have more images of Sanders, but that the images were much more negatively biased. Overall 6 of the 7 media outlets investigated, had more images of Clinton than Sanders associated with 'Joy'. CNN in particular did not have single photo of Sanders associated with 'Joy', but they had two photos of him associated with 'Sorrow' (Kwak & An, 2016). In another development in the domain, Narwal et al. propose an automated assistant called UnbiasedCrowd to detect and prompt action on visual news bias (Narwal et al., 2017). The proposed tool is based on the idea that participants will review clusters of images to identify visual bias in news stories. An online bot is then designed to engage with users of social media exposed to these images to prompt action. The system is in ongoing development and has recently undergone a pilot study.

2.3.6.1.6 Position Bias in News on the WWW

Position bias has long been studied in search, where it has been shown that users are heavily influenced by the position in the rankings of returned search results (Craswell et al., 2008). Users are more likely to consider results at the top as being more relevant, important, and likely to satisfy their information needs. Recently position bias or prominence bias has begun to be looked at in the dissemination of news online.

The homepages of news websites have increasingly more in common with search engine results pages than they do with the front pages of a print newspaper. The front pages of newspapers form a static snapshot of the day's events. In comparison, the homepages of news websites are constantly changing, due to changing position or prominence of articles, article churn, streaming news services, trending content, personalisation, and increased social media integration. They are also increasingly a unique and unrepeatable experience. The aforementioned study by Chakraborty et al. demonstrated that trending news can create temporal coverage bias in online news (Chakraborty et al., 2015). This was due to the position of news articles in the trending news section of the website. This demonstrates the effect of promoting or highlighting popular news which may not be as important to the consumer and the resulting detrimental effect of burying boring news which may be more important. Yet there are insufficient studies on the impact of changing the size

position or prominence of article headlines and blurbs, article churn, trending content, personalisation, and increased social media integration, or the combination of some or all of these, on the consumer.

Ahmed et al. note and attempted to tackle some of these issues with their ‘fair and balanced’ method to present news stories (Ahmed et al., 2012). While their contribution focuses on a streaming news service rather than a homepage, they identify and tackle many of the issues, such as relevance, recency and personalisation, which are similar to the issues on a homepage. They also highlight a potential solution to position bias in the streaming service. In another relevant contribution, Lagun and Lalmas also highlight a strong positional bias in news article reading (Lagun & Lalmas, 2016). Their study shows that content at the top of a news homepage or an article page, will receive significantly more attention than content at the bottom of the page. Though those who demonstrate an interest in the article will spend quite a lot of time at the bottom of the page. Ma et al. also demonstrate the impact of position bias on user fatigue in the Bing Now¹⁵ news recommendation service (Ma et al., 2016). Despite the work on position bias in search and in news streaming webpages, no comprehensive work has been done on position or prominence bias of news articles on news homepages. This is important, as when users visit a news agencies homepage, they gain an initial impression of the relative importance of the day’s events based on the position, prominence, coverage, headlines and density of content on a topic. Despite this, there have been no studies on position or prominence bias including, article churn, streaming news services, trending content, personalisation, increased social media integration, layout, density, and overall coverage of different stories, on the consumers. It must be noted that position bias may also not be the most appropriate term considering its use in search and lately in news streaming services. Consequently, prominence bias is considered by this thesis as a more appropriate term.

2.3.6.1.7 *Design Bias in News on the WWW*

Design bias is an arch term for all the aforementioned visual biases, though it is rarely used in the literature. It is however used by the Society for News Design (SND)¹⁶ who caution against visual design biases such as the deliberate or accidental imbalances in coverage. This includes; prominence, headlines, imagery, or graphics. The SND is also regularly referenced in literature, often when referring to professional codes of ethics or datasets (Diddi et al., 2014; Lowrey, 2002; Zeldes et al., 2008). As part of their outreach and training, the SND Foundation¹⁷ recently held a design panel in Northwestern University titled “How designers can try to correct for the unconscious bias in design”.¹⁸ The panel highlighted that bias in

¹⁵ <https://www.bing.com/news>

¹⁶ <https://www.snd.org/> The Society for News Design (SND) was set up in 1979 for visual news media professionals who create print, web or mobile news publications.

¹⁷ <https://www.snd.org/foundation/> The Society for News Design Foundation is involved in outreach, the provision of training grants and scholarships and the provision of research grants.

¹⁸ Northwestern University’s Knight Lab is a community of designers, developers, students and educators working on design and visual experiments in journalism including the building of opensource tools for the news community to visually tell stories. The SND panel was hosted on <https://knightlab.northwestern.edu/2016/04/15/snd-panel-even-savvy-designers-can-introduce-bias-into-design/>

designer's output is the result of five influences; past experience, community influence, indirect messaging, direct messaging, and reinforcement. They maintain that news designers also hold particular influence over two of the ways bias is formed in the receiver of the message, day-to-day messaging and reinforcement.

In summary, this section has highlighted some of the main forms of bias that have been studied in the dissemination of news online including; headline, coverage, photographic, presentation and visual. It also highlighted the relative lack of research into position or prominence bias. A broader overview of the domain and relevant studies is available in Table 2-3. Most apparent here is the complete lack of research on the impact of news website design on the perception of bias. This extensive literature review was unable to find a single relevant empirical investigation that has been undertaken on the subject. This is despite its obvious importance to the means of delivery.

2.3.7 Bias Frameworks: Theoretical

To measure the impact of news website design on perceived bias, a search was undertaken for any models, theories or frameworks to explain how users identify and form judgements of bias, specifically online.

Despite the dearth of published underlying theory on bias encountered during the literature search, it was initially felt that an in-depth search would reveal models, theories and frameworks explaining how individuals recognise or form judgements of bias in the news they consume. However, none was identified for any medium of dissemination of news, print, radio, television, or online. In fact, there is very little theoretical work on bias outside of economics, such as the work of Tversky and Kahneman who investigated heuristics and decision biases under judgements of uncertainty (Tversky & Kahneman, 1974), and political communication, such as Zaller's two message model of political communication (J. R. Zaller, 1992).

This revelation was in direct comparison to the ten models, theories and frameworks identified explaining how users judge the credibility of information online, which were identified and are presented later in this chapter, see Table 2-1. Bias, along with trustworthiness, expertise, accuracy, fairness, and depth of coverage among others, is a core dimension and measure of credibility.

2.3.8 Bias Frameworks: Experimentation

Forewarned by the disappointment in being unable to identify any theoretical frameworks to explain how individuals identify or form judgements of bias online, or in any other medium for that matter, it came as no surprise that there were also no frameworks for undertaking experiments on bias online. There are also no empirical frameworks for any other medium of dissemination. This is despite the myriad of studies conducted over many years presented heretofore.

Consequently, an exhaustive yet ultimately futile search was undertaken for empirical frameworks in the arch domain of credibility, see section 2.2.13. Despite identifying ten models, theories and frameworks providing underlying theories as to how users form judgements of credibility online, none were found for experimentation. Thus, the decision was made to develop a framework that could be used to design experiments to measure credibility, or any of its individual dimensions and measures. This was based on the requirements identified in section 2.2.12.

2.3.9 A Lack of Frameworks for Identifying, Measuring, or Eliminating Bias in the News

A thorough search of the literature using combinations of terms related to “Bias” and “Framework” revealed no theoretical or empirical frameworks relating to bias and the news similar to the work proposed in this thesis. The few examples found include a Master’s thesis on the impact of ownership on objectivity in news organisations (X. Wang, 2003), a tool for identifying and eliminating bias in social science (Burke & Eichler, 2006), Park’s contribution for aspect level browsing to reduce bias (Park et al., 2009), and an investigation of partisan bias in television news (Diddi et al., 2014).

2.3.10 Deductions from the State of the Art on Bias

There are several evident deductions which can be drawn from this literature review. 1) There is a distinct lack of underlying theoretical research on bias. 2) Bias in the dissemination of news has been extensively studied in traditional mediums, especially in print newspapers. 3) Bias has, to a much lesser extent, been studied in the dissemination of news in the WWW. 4) This literature review found no relevant work on the impact of news website design or webpage aesthetics on perceived bias. This is despite the fact, that a websites’ design, including its layout, colour scheme, and branding, are among its most obvious attributes and have been extensively studied as to their impact on other dimensions of credibility such as trust and reputation. 5) There is a dearth of underlying models, theories, or frameworks, explaining how users form judgements of bias online. 6) There are also no empirical frameworks to undertake such research.

2.3.11 Experiments on the Impact of News Website Design on Perceived Bias

To satisfy the research objectives, and answer the research question, the following course of action was decided. First, an empirical framework would be developed to ascertain the impact of website design or webpage aesthetics on credibility, and any of its dimensions and measures such as bias. This would form one of the minor contributions of this thesis. Second, experiments would be undertaken on the impact of a website’s design and/or webpage aesthetics on the perception of bias. The first experiment would investigate the impact of individual features, such as advertising, social media integration, or user comment facilities. The second experiment would investigate wider characteristics, such as the visual quality or the level of professionalism. The third experiment would build on the findings of experiments one and two, and investigate, through distortion, if it is possible to predictively impact perceived bias in the content of a news article.

The fundamental supposition behind undertaking this research, is that users, at least initially, heuristically evaluate news on the WWW. This is not completely a new idea. Baum and Gussin previously pointed out the role of heuristics in judgements of bias in online news: *“Our results suggest that in an increasingly fragmented media marketplace, individuals not only distinguish between media outlets but, more importantly, outlet brand names, and the reputations they carry, function as heuristics, heavily influencing perceptions of bias in content.”* (M. A. Baum & Gussin, 2008). While this research is not designed to investigate brand name or reputation, the supposition behind the overall research hypotheses is that users pick up on small design and aesthetic cues which they use to make initial judgements about the message.

2.3.12 Conclusions on Bias

The breadth and depth of research on bias has demonstrated that the majority of experiments in the domain do not focus on the textual, spoken, or visual content of newspapers, radio and television, but rather focus on the surrounding attributes and the means of dissemination which may influence the message. This review has demonstrated that no serious work has been undertaken on the possible impact the visual presentation of news on the WWW may have on perceived bias in news articles. The distinct lack of underlying theory to explain judgements of bias, and the complete lack of empirical frameworks for measuring perceived bias is also evident. This is contrasted with the abundance of both theoretical and empirical frameworks identified to explain and measure trust.

2.3.13 Summary on Bias

This State of the Art began with an introduction to bias described by Entman as: *“that curiously undertheorized staple of public discourse about the media.”* (Entman, 2007). The following sections provided further evidence of the lack of research despite the fact that the term has regularly been used as an accusation or invective to dismiss news or opinion for at least a century. A detailed description of the difficulties of measuring bias was then explored before historical tests for the construct from the literature is provided. The section then provided the most detailed classification of definitions existent in the literature including a categorisation of definitions by related terms. An overview of bias in the three main stages of the news pipeline; production, dissemination, and consumption were then provided to demonstrate an understanding of bias in the overall domain. This included the first classification of global, domain, organisational, and journalistic and editorial biases that impact the production of news.

The State of the Art then continued with a detailed classification of biases which impact the dissemination of news in the traditional mediums of radio, print, and television. This serves to emphasise the dearth of comparable research on bias on the WWW. It quickly becomes apparent that no serious research has been done on the impact of the visual presentation of a news on the WWW on the perception of bias in news articles, despite it being one the most conspicuous possible influences. This research has thus set forth to

remedy this situation. To further demonstrate this, a small selection of relevant studies on some of the individual biases affecting each of the traditional domains are then discussed in detail. In comparison, only a small amount of studies investigating bias on the WWW could be identified. Lastly, this section of the State of the Art highlighted the lack of theoretical or empirical models, theories, or frameworks explaining how judgements of bias are made or how it should be measured.

2.4 Trust

Trust, along with expertise are two of the earliest defined dimensions of credibility (Hovland et al., 1953, p. 21). This section will begin by detailing relevant studies on the impact of website design on perceived trust which is also a dimension and measure of credibility. It also highlights ten theoretical models, and five theoretical frameworks explaining how users form judgements of trust online, further highlighting the lack of underpinning theory for bias and the disparity between the two fields. It also highlights six empirical frameworks put forth to measure the construct online, again showing the disparity in the domains. The aim of this section is threefold. First, to show that the impact of website design has been looked at among other dimensions and measures of credibility. Second, to highlight any models, theories or frameworks that have been put forward to explain how users judge trust online. Third, to investigate whether any empirical frameworks for measuring trust online exist, and if so, whether they might be of any use to measuring the construct of bias.

2.4.1 Introduction

A general definition of trust as a peer to peer construct from Mayer is: *“the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the truster, irrespective of the ability to monitor or control that other party”* (Mayer et al., 1995). Trust has been defined by Shneiderman in an online business to consumer context in his contribution ‘Designing Trust into Online Experiences’ as: *“The positive expectation a person has for another person or an organization based on past performance and truthful guarantees.”* (Shneiderman, 2000). Further definitions and in depth analysis of the construct as it pertains to WWW applications and ecommerce is provided by (Grandison & Sloman, 2000; Yao-Hua Tan, 2000).

Trust, even more than bias, is considered among the core dimensions of credibility. In fact, Hovland and his communications and change group at Yale, defined Trust and Expertise as the primary attributes of a credible source. They maintain that the credibility given to the arguments of a source is the resulting permutation of these two dimensions. *“In any given case, the weight given a communicator’s assertions by the audience will depend upon both of these factors and the resultant value can be referred to as the ‘credibility’ of the communicator.”* (Hovland et al., 1953, p. 21). As previously stated, credibility is a multi-dimensional construct, however measures such as trust, expertise, believability, accuracy, bias, fairness

and completeness of information have regularly been used in studies to measure it, especially online (Bucy, 2003; Flanagin & Metzger, 2007; Johnson & Kaye, 1998).

2.4.2 *The Impact of Website Design on Trust*

The impact of website design on the perception of trust has received a lot of attention, specifically in the areas of ecommerce and e-health. Trust has long been considered an important factor for websites representing business for which trust was also important, such as banks, legal professions, health care providers, and government organisations. However, it was the advent of ecommerce that brought it to the fore as a design consideration.

Design is considered one of the core dimensions of ‘website quality’ which is used to measure the performance of websites. Website quality is a multidimensional construct, with the number of dimensions varying between three (Rocha, 2012) and as many as twelve (S. Kim & Stoel, 2004a). Often the type of website or its domain impacts the number of dimensions, with websites that sell a product or service typically having more due to the increased interactivity and services they offer, and the resulting increased scrutiny they are under. Dimensions of website quality include; appearance, information quality, service entertainment, fit-to-task, response time, security, interactivity, intuitiveness, and privacy, among others (Barnes & Vidgen, 2000; S. Kim & Stoel, 2004a; H.-F. Lin, 2007; Loiacono et al., 2002). Many of these dimensions are the overarching terms resulting from factor analysis of user interviews or surveys. There are large areas of overlap between them and the exact terminology can change from study to study. Though many use the term design, such as Lin (H.-F. Lin, 2007), it may also be known as visual appeal, appearance, visual design, graphic design, visual interface, or aesthetic treatment. The relationship of design to website quality is analogous to the relationship of bias and trust to credibility. Both website quality and credibility are multidimensional constructs, the dimensions of which vary depending on the situation. It should also be noted that trust has also been identified as a dimension of website quality (Hyejeong Kim & Niehm, 2009; S. Kim & Stoel, 2004a, 2004b; H.-F. Lin, 2007; Loiacono et al., 2002; Wolfenbarger & Gilly, 2003). In fact, the perceived quality of a website is: *“highly correlated with trusting beliefs and intentions”* (McKnight et al., 2002).

Although good website design is synonymous with usability, for brevity, the following brief review only highlights contributions that focus on the design, visual quality, or aesthetics.

A plethora of empirical studies have looked at the impact of website design on trust. A study by Roy et al. was among the first to highlight a strong relationship between interface quality and user trust (Roy et al., 2001). Gefen and Straub investigated the impact of social presence, typified by pictures or personal details of staff, on trust in Business to Consumer (B2C) websites. Their study demonstrated that increased social presence increased trust, which in turn increased purchase intention (Gefen & Straub, 2003). In another

study on the impact of dimensions of website quality, including the impact of website design and trust, on customer satisfaction, Lin demonstrated that both were significant determinants of customer satisfaction (H.-F. Lin, 2007). Ganguly et al. also demonstrated that website design positively influences trust, thus increasing purchase intention (Ganguly et al., 2010). In a departure from many in the domain, Ou and Sia maintain that trust and distrust are separate concepts by reason of their distinct cognition (Ou & Sia, 2010). Their contribution includes a framework for testing the impact of specific features of a website's design on trust and distrust in the user. The study looked at eleven website attributes divided into four categories, however no conclusive findings regarding the impact of specific features were evident. Seckler et al. also investigated the impact of specific characteristics of websites on trust and distrust (Seckler et al., 2015). Their study revealed that 'graphic design' was frequently mentioned in experiences which resulted in distrust, but not in experiences of trust. A similar finding was also found for structure design.

The impact of specific features or characteristics of a website's design on trust have also been investigated. The following paragraphs present a small proportion of the research on logos, colour schemes, and photos.

The impact of website logos on perception of trust has been investigated by Lowry et al. who utilised source credibility theory to design and develop logos which communicate credibility, including expertise and trustworthiness (P. B. Lowry et al., 2014). They also show that the positive impact on perceptions of trust in the user are magnified when the design of the website complements and extends the design. It should be noted that Lowry's work is heavily relied on later in the following credibility section of this thesis. This study also focuses on two other core dimensions of credibility, expertise and trustworthiness.

The impact of colours in website design on perceptions of trust has also been studied extensively. Kim and Moon demonstrated that through the manipulation of design factors, including colour, it was possible to influence perceptions of trust in cyber banking (J. Kim & Moon, 1998). Based on their earlier study, which demonstrated that visual design, including the colour, impacted trust, satisfaction and e-loyalty, Cyr et al. specifically investigated the impact of colour appeal, again on trust, satisfaction and e-loyalty (Cyr, 2008; Cyr et al., 2010). Their eye tracking studies demonstrated that colour is a significant determinant for website trust. In a more in depth study, Pelet and Papadopoulou found that colours that are low in saturation and brightness infuse trust, while colours that vivid and are high in saturation have a negative impact on trust (Pelet & Papadopoulou, 2012).

The impact of photos on trust in ecommerce websites has also been regularly investigated. Steinbrück et al. found that including photos of customer service agents had a significant positive impact on perceptions of trust in online banking websites (Steinbrück et al., 2002). Riegelsberger et al. also found that: "*the trustworthiness of low-trust sites can be boosted by adding a photo*". However, when averaged across websites, they found no significant impact, and if anything found that they had a levelling effect

(Riegelsberger et al., 2003). More recently Bente et al. demonstrated that trustworthy photos have a positive impact on potential buyers trust (Bente et al., 2012). Surprisingly, they also found that missing photos and information was worse than untrustworthy photos. Similar results were also found by Ert et al. in their study of photos on trustworthiness in the sharing economy (Ert et al., 2016).

The highlighted papers presented here are but a fraction of the literature on the subject. For a more in depth accounting of the impact of website design dimensions on trust, Karimov et al. present an exemplary synthesis of the literature in the domain (Karimov et al., 2011).

2.4.3 Theoretical and Empirical Models, Theories and Frameworks for Trust

Like bias, trust is a perceived value placed upon a person, message, or resource, as to the quality of the information they provide. However, unlike bias, there is a wide range of studies on the impact of website design and aesthetics on the perception of trust, a small proportion of which are presented above, see section 2.4.2. Also, unlike bias, there have been significant advances in the theoretical understanding of trust on the WWW, including its measurement. Many of the most prominent researchers in the domain including Lucassen et al., who's Revised 3S Models of Credibility Evaluation is reviewed in section 2.2.8, also explicitly link trust and credibility. Consequently, it was decided to review the domain of trust to see if any theoretical or empirical frameworks could be discovered which would inform how judgements of bias might be made or how bias should be measured online and whether they could be used in this research.

This review identified ten theoretical models and five theoretical frameworks explaining how users from judgements of trust on the WWW. It also identified six empirical frameworks for measuring trust online. For brevity, these were moved to sections 9.8.3 and 9.8.4. For a range of reasons, these were not suitable for measuring bias. A conclusion and summary of the findings on trust are also presented in sections 9.8.5 and 9.8.6 respectively.

2.5 Conclusions on the State of the Art

The main conclusions that can be drawn from this State of the Art are as following. 1) There is very little underpinning theory on bias. 2) Bias is a subjective construct, and the myriad of definitions and the range of research in the domain reflects this. 3) It has been studied at almost every stage of the news pipeline, production, dissemination, and consumption, across all mediums. However, there is a distinct lack of research on the impact of bias during the dissemination of news on the WWW. 4) There are no models, theories, or frameworks explaining how users arrive at, or form judgements of bias online, or in any other medium, nor are there any empirical frameworks for undertaking such research. 5) In comparison, there is almost a cornucopia of such in the aligned domain of trust, however the differences between the research application areas and the fundamental differences between the constructs means that they were only useful for informing the process. 6) Bias is a core dimension and measure of credibility, especially when judging

news on the WWW. 7) There are a range of studies demonstrating the impact of website design and webpage aesthetics on perceived credibility, yet none investigating the impact of news website design or webpage aesthetics on bias. 8) There are fully ten MTFs, explaining how users form judgements of credibility online, yet there is no empirical framework. 9) The development of such a framework for undertaking experiments on credibility or on any of its dimensions and measures such as bias would be an important contribution to the domain.

2.6 Summary on State of the Art

This State of the Art began by exploring credibility, the overarching research domain of this thesis. Definitions and key terms were provided and clarified before ten human-centric MTFs put forth to explain how judgements of credibility are made online were reviewed. The main underlying theories for these, namely the ELM, HSM and the C&AHIP were also reviewed which maintain that users may make judgements of information by the central route, systematic approach, or controlled processes, or the peripheral route, heuristic strategy, or automatic processes. Of each respectively. A broad range of studies were then reviewed which demonstrate the impact of website design on perceptions of credibility online.

The specific research domain of bias, a core dimension and measure of credibility was then explored. It provided an in-depth analysis including a historical overview of research, definitions, and relevant studies. It demonstrated that there is a dearth empirical study on the impact of news website design and webpage aesthetics on the perception of bias in the articles they contain. This is in direct comparison to the broad range of research investigating bias at every stage of the news production pipeline, and in almost every conceivable means of dissemination in traditional media. The State of the Art also demonstrated a lack of underlying human-centric and empirical, models, theories, and frameworks as to how judgements of bias are made.

To demonstrate this fact, the aligned research area of trust, another core dimension and measure of credibility was explored. It revealed an extensive collection of research demonstrating the impact of website design and webpage aesthetics on trust, and a broad range of theoretical and empirical frameworks.

While the focus of this research is on bias, the opportunity was recognised to develop an empirical framework for measuring the impact of website design on credibility, and on any of its dimensions and measures such as bias as this would have more utility in the domain. In summation, two facts are apparent. First, research needs to be undertaken on the impact of news website's design, reflected in webpage's aesthetic, on the perception of bias in the news articles they contain. Second, a comprehensive research framework and experiment platform needs to be developed for such. This State of the Art satisfied the first research objective highlighted in section 1.2.2, namely; Research the State of the Art on measuring bias and what frameworks are used.

3 CAFE - An Experiment Framework and Platform for the Measurement of Perceived Credibility, or any of its Dimensions or Measures

3.1 Introduction

This chapter will introduce the Credibility Assessment Framework for Experimentation (CAFE), a structured process and accompanying experiment platform to design, develop, and deploy experiments for measuring credibility online, or for any of its dimensions or measures. Researchers using the framework engage in a structured and iterative process to define each phase, stage, and option they should consider for their experiment. Thus, the result of following the framework process can be used as an experiment plan. The CAFE framework is accompanied by an experiment platform which was used extensively in the execution of the experiments conducted as part of this thesis.

3.2 Motivation

Overall the majority of research published in the domain is of a high quality. It is likely that many of the experienced researchers behind these studies have already adopted similar experiment design processes, though they may not be as explicit, formalised, structured, or as comprehensive as CAFE. There are also many examples of low-quality studies with one or more issues. Those familiar with the academic review process for conference and journal papers will be aware of the many studies which fail to get published due to issues with experiment design, execution, statistical analysis etc. Many of these issues and the resulting rejections could have been prevented with appropriate planning and design. However, there is no published experiment design framework for novice researchers or those new to the domain to help them design, develop, and deploy an experiment. It is hoped that CAFE may help others avoid such pitfalls in future. The lack of a published framework has also meant that no baseline or formal process exists to which researchers can easily refer to, or which they can make stated deviations from. These issues have impacted experiment rigor, reproducibility, and have made comparisons of different studies investigating the same question or issue difficult, e.g. comparing multiple studies investigating the credibility of different mediums.

3.3 The CAFE Framework

CAFE is an open empirical framework and platform designed to support the design and deployment of complex experiments measuring credibility online, and its dimensions or measures. The CAFE framework and accompanying experiment platform aim to help researchers to design and deploy a broad range of studies. Examples include studies investigating specific individual features or characteristics of a website's

design, comparative evaluations, or investigations of the construct of credibility itself. The framework is both theory and instrument. It has been successfully used to measure the impact of news website presentation on the perception of bias, a core dimension and measure of credibility. These experiments are detailed in chapter 4, chapter 5, and chapter 6 of this thesis. All three have been published as full conference papers (Spillane, Lawless, et al., 2020; Spillane et al., 2017a, 2017b, 2018).

The framework and platform have also been used successfully by three students in Trinity College to complete of a final year undergraduate project and two Master's research projects. These projects were guided by the author of this thesis.

A 4th year computer science student, used CAFE and the experiment platform to design and deploy an experiment entitled: "Kicking Them While They're Down". This project investigated whether there was favourable / unfavourable bias in photographs of four Premiership football managers in three UK publications when their team was successful / unsuccessful in the Premier League. It demonstrated that there is a correlation between the results of the match and the favourability of the photographs used by the three publications.

The first Masters student used the CAFE framework and experiment platform to design and deeply an experiment entitled: "Can News Graphics Impact the Perception of Bias". It tested the influence of different types of information graphics on the perception of bias in the news article. It found that the information graphics did not influence the perception of bias.

The second Masters student used the framework and platform to measure the impact of increasing the visual 'Tabloidization' of the homepages of four existing news websites on initial impressions of credibility. A tabloid website was also included as a baseline. Tabloidization was achieved by incrementally and commonly adapting individual elements of each homepage such as headlines, colour schemes, and the content of images, in stages to increasingly reflect a Tabloid style. To measure initial impressions, participants were exposed to each distorted homepage for a short period of time. This has since been published at CHI 2020 (Spillane, Hoe, et al., 2020).

The framework and accompanying experiment platform are designed to be used by researchers, particularly novice researchers or those new to the domain, as an aid for deploying similar experiments, and by the wider, non-academic community, particularly those who may be involved in the design and development of websites in domains where credibility is a singular concern such as news, ecommerce and healthcare. Novice researchers will benefit more than experienced researchers as it will help them to avoid many of the pitfalls and mistakes such as: choosing the most appropriate measures, scales, and statistical tests, avoiding underpowered experiments, and timely selection of appropriate target venues for publication of

their research. The framework and platform also help to solve three major issues; reproducibility, longitudinal comparisons, ease of deployment.

3.3.1 An Open Framework

By ‘open’ framework, CAFE is designed to be adapted by researchers to their experiment aims, research objectives, and other experiment influences such as the research domain. It is a structured approach or process for researchers to follow to design and deploy experiments to measure credibility, or any of its dimensions and measures. The result of following the framework is a defined experiment plan for each phase, sub-stage, and option which has to be considered to increase the likelihood of successfully deploying an experiment.

3.3.2 Framework Description, Objectives, Use, and Purpose

The following describes the main objectives and purpose of using the CAFE framework. Figure 3-1 depicts a hypothetical credibility experiment with some of the possible phases, stages, and considerations a researcher has noted. It is divided into four rows: Phase, Stage, Consideration, and Impact. Each of the phases, Design, Build, Initiation, Execution, and Analysis, represent some of the main phases that a typical experiment to measure credibility goes through. It should be noted that the framework is designed to help a researcher map out or flesh out an idea into an experiment plan, it is not designed to help them generate an experiment idea. The design of CAFE is based on the requirements which were identified in 2.2.12.

3.3.2.1 Phases

To utilise the framework, the researcher begins with a blank page of paper, whiteboard, spreadsheet, word document etc, whichever they feel most comfortable working in. Often paper or whiteboards are the most useful for early stages before it is more formally presented in digital form. The researcher begins with the first row by noting down the main phases of their experiment lifecycle. They may choose to use the phases shown in the CAFE experiment framework shown in Figure 3-1. They may also add additional phases such as testing, reporting etc. as suits their individual purposes. The number and composition of the phases depends on how much of the experiment lifecycle the researcher wants to include in their experiment plan. The purpose of this part of the process is to get the researcher thinking about the experiment from start to finish, at an abstract level. Unlike the following stages or considerations, the phases defined by the researcher are unlikely to change much as the experiment plan develops. The five Phases: Design, Build, Initiation, Execution, and Analysis, depicted in Figure 3-1 represent the some of the main phases of a typical experiment to measure credibility or any of its dimensions and measures. Researchers may consider adding additional phases such as, pre-experiment content testing, platform testing, write up of experiment results, or reporting. As an open framework the researcher can decide how detailed they want their approach to be or how much of the experiment lifecycle they want to plan for.

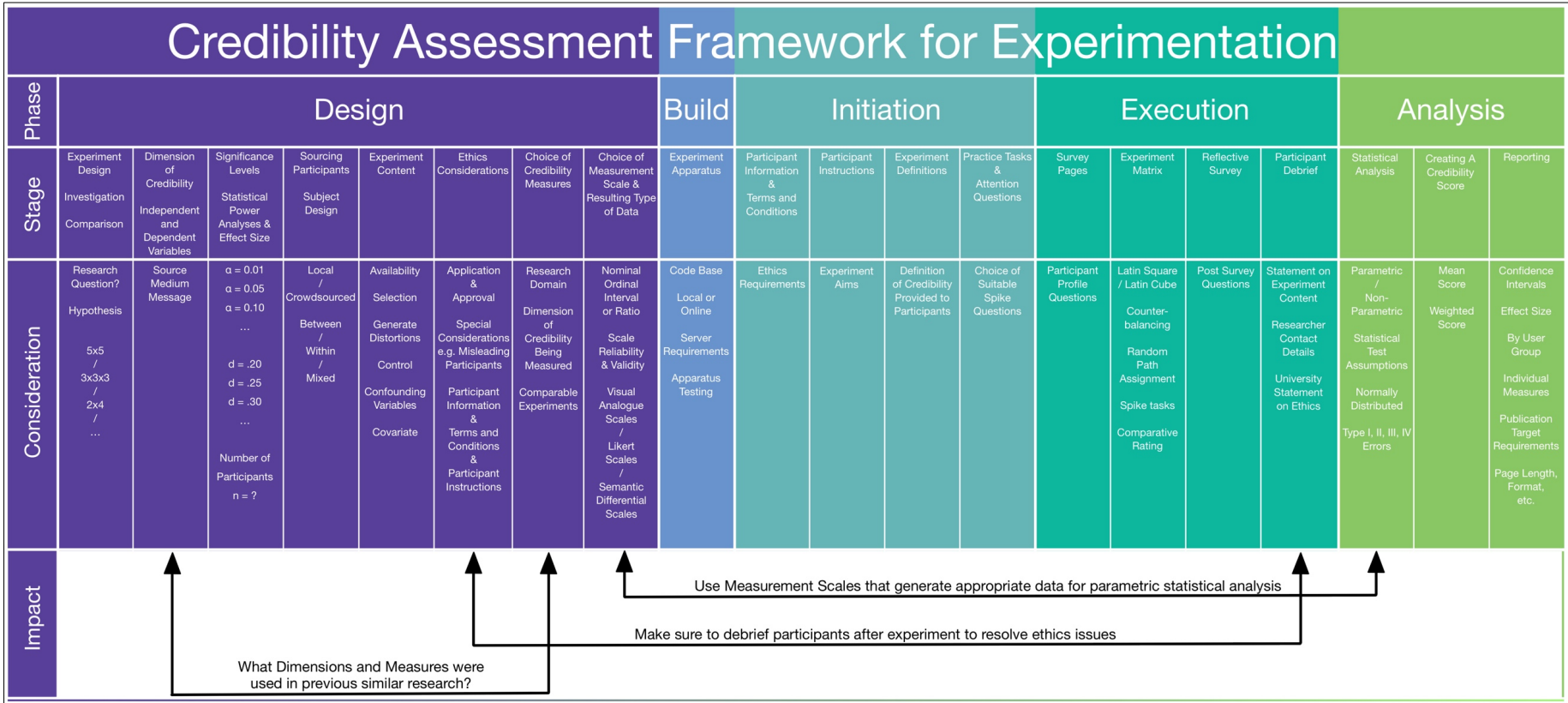


Figure 3-1 The Credibility Assessment Framework for Experimentation (CAFE).

3.3.2.2 *Stages*

The second part of the framework process involves the researcher adding a stages row under the phases row to their developing experimental plan. This is divided into columns to represent the individual stages, of each phase of the experiment, as can be seen in Figure 3-1. As an open framework, it is designed to be modified by the researcher. Thus, the combination of and the number of stages in each phase can be increased or decreased to suit the type of experiment, the experiment aim, and the level of detail the researcher wants to include. It should be noted that as the overall experiment begins to develop, and the complexities of the proposed research study become more apparent, the researcher may choose to add, remove, update or reconfigure the various stages they have defined under each phase. The CAFE framework is an iterative process, and this is to be expected. The stages included in Figure 3-1 are typical of many experiments in the domain and they may help other researchers initially to define some of the stages of their experiments. Figure 3-1 does not include every stage, nor is it intended, or even possible to do so. A stage is indicative of a point in the experiment plan where a researcher feels they need to make a decision between multiple considerations. These are discussed below in section 3.3.2.3. Thus, the stages chosen by each researcher are particular to their experiments aims, the content being measured, and the type of experiment etc.

Beginning with the design phase, see Figure 3-1, the researcher begins by adding sub-stages, e.g. to defining the type of experiment being conducted and the dimension of credibility being measured. The early stages under the experiment design phase tend to focus on the larger or more important decisions. The researcher adds additional stages under each phase depending on the level of detail they want to include. Common stages in the design phase of an experiment that a researcher may want to define include: the experiment design (3x3, 4x2 etc.), defining the dimension of credibility being measured (source, medium or message), desired Statistical Power (.20, .30 etc.), source of participants (crowdsourcing, student samples, lab group etc.), sourcing experiment content (existing websites, existing news articles, or creating new websites etc.), ethics considerations (misleading or lying to participants, consent forms, debriefing), choice of measurement scales (Likert, VAS, Semantic Differential etc.). Figure 3-1 also displays some common stages in the Build, Initiation, Execution, and Analysis phases which may help other researchers to define their own.

The combination of stages included by a researcher in their framework will depend greatly on the particular experiment aims, the type of experiment being conducted, and the content being assessed, etc. It should also be noted that the decisions made at each stage will also affect the choice of and combination of other stages in the experiment design. The CAFE framework shown in Figure 3-1 cannot list every possible experiment stage. It is a structured process by which the researcher maps out their own unique experiment design, to create a similar representation.

It should be noted that while modifying the open framework, researchers need to ensure the accuracy and integrity of the framework as a representation of the intended assessment. The main method of achieving this, if the framework is used correctly, is to consider the impact of the decision that they make under each stage that they include in their version on earlier decisions that made, and upcoming decisions to be made under future phases and stages. Researchers should always be cognisant of the fact that each decision has implications, e.g. the type of data collected impacts the type of statistical analysis, or the statistical power may impact which venue will accept a resulting publication. Thus, using the framework is meant to be an iterative process. Of course, for some researchers, especially novice researchers or those new to the domain, the implications may not be clear. Therefore, they are encouraged to show this to supervising researchers or those with more experience to get their input. In this way, the framework serves as a communication tool, and as a means of planning experiments.

3.3.2.3 *Considerations*

Considerations are the options the researcher must consider or decide between at each stage of their experiment design process. The considerations row in Figure 3-1 depicts some of the options for each of the stages shown in the row above. For some common stages which may be specified by the researcher, e.g. the dimension of credibility being measured (source, medium, or message), or the type of design (between subjects, within subjects, mixed), the number of options the researcher must consider are limited, and often relatively static. However, for other stages such as the desired Statistical Power, or the number of participants required, the considerations are much more numerous, and they are constantly changing depending on a number of the other stages and considerations. Considerations such as experiment design (2x3, 3x4, etc.), desired Statistical Power, the number of participants that can be sourced, or an experiment's budget, all influence each other. E.g. increasing the size or complexity of an experiment design from a 2x3 to a 3x4 will reduce Statistical Power, unless there is an increase in participant numbers. Thus, considering different options or under one stage can impact options at another stage.

The purpose of the considerations row is to make the researcher explicitly consider and record the various options specific to their evolving experiment. This helps researchers to consider alternative options and how they might affect other stages of their experiment. The choice of phases and the number of sub-stages for each phase that the researcher chooses to include, means that there is a huge range of possible considerations for each experiment. Thus, in Figure 3-1, only a small proportion of considerations specific to one hypothetical experiment design are shown.

3.3.2.4 *Impact*

As noted previously, many of the options a researcher must consider impact other stages and considerations in a proposed experiment. E.g. which dimension of credibility being measured, source, medium, or

message, will impact which dimensions of the construct of credibility a researcher will chose to conceptualise the construct e.g. trustworthiness, expertise, goodwill etc. While the experiment design and the measurement scales used in the experiment will impact the statistical tests which can be used. Figure 3-1 shows three example impact points on the bottom row where the researcher has identified points of contention or made a note to improve the experiment design.

The purpose of the impact row is to enable researchers to note what other stages and considerations are affected by a choice between options at another stage in the design process. Of course, it is only possible to note incompatibilities or likely issues if they are known. This row also enables other researchers in a team, e.g. PhD supervisors, or more experienced researchers consulted on an experiment, to note any issues they foresee in an experiment design which need to be resolved. Thus, it helps to enforce experiment rigor.

3.3.3 Related Frameworks

The CAFE framework and platform is designed to aid researchers in the design and execution of experiments measuring credibility or any of its dimensions and measures. However, just as good design and usability are closely related but separate aspects of a good website, credibility should not always be thought of in its own. Though out of scope for this thesis, readers may be interested in some related frameworks that measure values not directly related to credibility or any of its dimensions and measures, but which it may be beneficial to be cognisant of. These include web quality frameworks, (Loiacono et al., 2007; Rocha, 2012; Wolfenbarger & Gilly, 2003), web experience frameworks (Mohd-Any et al., 2015; Sauro, 2015; Wolfenbarger & Gilly, 2002), and usability frameworks (Dingli & Cassar, 2014; Dingli & Mifsud, 2011; Nazir et al., 2018), and the trust frameworks discussed in detail in section 2.4.3. Some of these also include AB testing platforms. It may even be possible to use the accompanying experiment platform to undertake experiments supported by these frameworks.

3.4 Benefits of Using the CAFE Framework and Experiment Platform

There are numerous benefits to using CAFE, including:

- It supports experimental rigour
- It can act as evidence of a planned investigation
- It helps researchers to identify potential pitfalls, conflicts, or issues before they arise
- It helps to make sure that each member of a team's understanding is in unison
- The experiment platform will enable others, especially those without advanced technical capabilities, to deploy experiments with greater ease in future
- It can help with working out experiment requirements and timelines
- It focuses an experiment on a research aim and prevent research drift

- It supports preregistration hypothesis on platforms such as Open Science Framework (OSF)¹⁹
- It aids in the publication of Fair, Accessible, Interoperable, and Reusable (FAIR) data publishing to improve repeatability in scientific experiments (M. D. Wilkinson et al., 2016)

3.5 *An Accompanying Experiment Platform*

As part of the development and testing of the CAFE framework, an experiment platform with accompanying documentation was also developed. The first version of this is now publicly available²⁰.

3.5.1 *Supported Experiment Designs*

In its current first iteration, it supports investigative experiments on the impact of individual features and characteristics of a website's design, reflected in a webpage's aesthetic on perceived credibility, or any of its dimensions and measures. Essentially these can be described as treatment and effect experiments. The treatment may be the inclusion, omission or distortions of a single or multiple elements, and the effect is the change in ratings for credibility. It can also be used to design and deploy experiments comparing the websites or webpages of different organisations, or different versions of the same website or webpage belonging to one organisation. These are often described as A/B or A/B type experiments. They may be run as side by side comparisons or shown to users individually. With careful and considered deployment it could also be used to measure the impact of variations in message on the perception of credibility. This would require some slight modification to the code base so that it displays text rather than images of webpages. It could also be easily modified to investigate credibility in other forms of message and medium such as video or print.

3.5.2 *Description of Default Functionality*

The experiment platform is capable of being used to run complex end-to-end experiments from the outset when it is successfully deployed using its default setting in the documentation provided. The default end-to-end experiment has five stages.

Stage one focuses on setup. It includes the basic landing pages the participant sees when they first decide to partake in the experiment. Note, the platform is designed for online deployment, so that participants may conduct the experiment from any location after clicking on a link either sent directly to them, or which is advertised in a crowdsourcing website. However, the platform may also be deployed locally on a single computer and participants asked to undertake the experiment in controlled conditions. The setup stage also includes additional pages which the researcher may use for tasks such as gaining informed consent, or describing the experiment in detail.

¹⁹ <https://osf.io/>

²⁰ Version 1.0 of the experiment platform may be found at: https://github.com/brendantcd/Credibility_experiment_apparatus.git

Stage two focuses on the provision of participant instructions and/or training such as dry runs of the experiment tasks so that the participant can familiarise themselves with the interface. A research could also include instructional videos of what the participant is meant to do during this stage if they required.

Stage three is the experiment proper. In its current default form, the platform supports a within subject incomplete counterbalanced measures design of size $N \times N$. This could be as small as a 2×2 or as large as 9×9 , or even larger. Incomplete counterbalancing is achieved by utilising a Latin square or cube, through which each participant is assigned at random to a diagonal path. Once a participant is randomly assigned to a path, the steps on that path are randomised. This is described in detail in section 3.6.5. Each step on a path is a new page, which contains an image of the webpage being rated, and the scales and measures the participant uses to rate that webpage. Participants are also given the opportunity to re-evaluate their ratings in side by side comparisons. At any stage, attention tasks may be inserted to check if the participant is paying attention to the task at hand.

Stage four focuses on reflective questions. These may be quantitative or qualitative in nature. The researcher may choose to use this opportunity to ask open questions about the experiment the participant just completed to gather additional information. All standard question formats are supported.

Stage five is for debriefing. As required by many ethics departments and research charters, participants should be debriefed at the end of the experiment and be given a final opportunity to opt out. In many cases the researcher may be required to take this opportunity to explain to participants what happened in the experiment, especially if e.g. fake news content was shown to the participant or if they were lied to for the purpose of the experiment. Researchers may also want or be required to include information such as their contact details.

3.5.3 *Development Technologies*

The platform has been built using mostly common open source technologies so that it can be easily maintained in future. The majority of the code is written in Python 2.7.3 using the Django Framework²¹ (Holovaty & Kaplan-Moss, 2009). This is a popular Python web development framework for the rapid deployment of websites. It is based on the Model Control View paradigm. It is deployed on a Linux Apache MySQL PHP (LAMP) stack²² (Lee & Ware, 2002), made public using Web Server Gateway Interface

²¹ <https://www.djangoproject.com/>

²² [https://en.wikipedia.org/wiki/LAMP_\(software_bundle\)](https://en.wikipedia.org/wiki/LAMP_(software_bundle))

(WSGI)²³ (Gardner, 2009), with a MySQL database²⁴ (DuBois, 2008). Some parts of the user interface such as pop up dialogue boxes and user rating scales are implemented in JavaScript²⁵.

3.5.4 *Extensibility*

One of the benefits of CAFE is that as it is designed to adapt to individual experiment needs, it is extensible and can adapt as new resources, experiment methodologies, statistical tests, or the focus of research change. The experiment platform is also extensible and will be updated to add new technologies and functionality for future experiments conducted outside the scope of this thesis. As additional functionality is added, the code will also be made available.

3.5.5 *Documentation*

The platform also comes with extensive documentation to help researchers deploy experiments. Currently the list of topics covered by the documentation includes:

- Setting up a new local project
- Deploying your project to a new server for the first time
- Installing MySQL on an Apache VM
- Connecting phpMyAdmin
- Modifying the code
- Testing the experiment locally and on the server
- Re-uploading – step by step
- Going live checklist
- Connecting it to crowdsourcing platforms
- Working with your experiment data
- Creating a baseline datafile
- Getting your data into SPSS

3.6 *Best Practices, Recommendations, and Novel Contributions to the Typical Credibility Experiment Lifecycle*

The vast majority of published research on credibility is of an excellent overall standard. However, those familiar with the review process have knowledge of the many studies which fail to get published due to issues with design, statistical analysis, or reporting. One of the main aims of CAFE is to help researchers identify potential issues early, so they can attend to them before they impair, severely degrade, or invalidate

²³ <https://www.python.org/dev/peps/pep-3333/>

²⁴ <https://www.mysql.com/>

²⁵ <https://www.javascript.com/>

an experiment and make it unpublishable, thus wasting valuable time and resources. Another way to improve the quality of an experiment is to ensure that it utilises best practices wherever possible. This section recommends existing or new best practices or introduces novel contributions to a typical experiment design. As many researchers already adhere to these best practices, they will likely be more beneficial to novice researchers or those new to the domain.

3.6.1 Statistical Power Analysis and Effect Size

Statistical power is the probability of detecting an effect, if the effect ‘is’ actually there, thus avoiding making a Type II error. These are discussed in detail in section 9.9 Since 1994 the American Psychological Association (APA) has recommended calculating and reporting Statistical Power and Effect Size (L. Wilkinson, 1999). Yet, many studies in the domain fail to do so. One of the most difficult decisions is deciding what constitutes a good Statistical Power and effect size. Based on the original work of Cohen, the analysis by Coe, recommendations of Nelson, and a look at comparative and aligned domains, the CAFE framework recommends a 95% Statistical Power of being able to detect an Effect Size of at least 0.25 when measuring credibility or any of its dimensions and measures, on news websites (Coe, 2002; J. Cohen, 1969; Nelson, 2013). Main Effects And Interactions. A detailed explanation of Statistical Power and Effect Size is provided in section 9.9.1 of the appendix and a detailed justification for this recommendation is provided in section 9.9 of the appendix.

3.6.2 Crowdsourcing Participants

The second existing best practice recommendation is for the use of crowdsourcing. It enables researchers to achieve strong Statistical Power, to gather a lot of data in a much shorter period than most existing methods, for a relatively small monetary investment. Arguments for crowdsourcing include the opportunity to significantly broaden the typical participant demographic, increasing Statistical Power, and speed of data collection. These arguments and others are detailed in section 9.10 of the appendix.

3.6.3 A Repository and Classification of Credibility Measures Used in Peer Reviewed Research

One of the core contributions of the CAFE framework and a significant and ongoing task, is the creation of an online repository and classification of measures of credibility. Due to its size and complexity, this is hosted online²⁶, but an overview of part of the repository is shown in Figure 3-2 with a closeup in Figure 3-3. It only includes measures that have been used in peer reviewed research to measure credibility online. Using this online repository, researchers can quickly ascertain which measures have been used most often, how they have been used, and in what context. This can differ significantly depending on whether the

²⁶ https://www.scss.tcd.ie/~spillab/Credibility_Measures/

researcher is attempting to measure or compare source, medium, or message, and the type of domain they are investigating e.g. news or health. This table forms the most comprehensive list of credibility measures in the literature and will serve as a reference point for future researchers.

3.6.3.1 *Description*

In the left column of Figure 3-2 and shown in more detail in Figure 3-3, are the individual measures, one per row. At the top, are some of the domains that credibility research is often undertaken in online, e.g. news, health, ecommerce, etc. These are further subdivided by the dimension of credibility being studied, usually source, message or medium. In the intersecting cells, are codes for each of the studies that the measure is used in, e.g. FM03 is the code for (Flanagin & Metzger, 2003). A key for each of these codes is also included in the repository, part of which is shown in Figure 3-5. As can be seen in the close up in Figure 3-3, some of the codes also have additional symbols beside them. These are designed to provide additional information such as whether the study was a comparison study, the type of scale used or what aspect of the website was looked at. A key for these codes is also shown in Figure 3-4.

3.6.3.2 *The Difficulty of Choosing Appropriate Measures*

Choosing appropriate dimensions and measures of credibility has proven to be a consistent issue in the domain. As Kioussis stated: *“Taken as a whole, it becomes clear that one overarching pattern resonating in credibility research is that scholars have failed to agree fully on the core dimensions of the concept”* (Kioussis, 2001b). Gaziano and McGrath also maintain that the choice of measures could influence perception (C. Gaziano & McGrath, 1986).

As shown previously in Figure 1-1, when researchers are conducting a credibility experiment, they first define which dimensions of credibility represent the construct as they encapsulate it, and as they want their participants to encapsulate it. As such they may end up with dimensions such as ‘fairness of the communicator to both sides of the argument’ or ‘concern for public good’ etc. The dimensions are dependent on the purpose of the experiment and what aspect of credibility the researcher is trying to measure and in what domain or context. While some dimensions are mainstays, there is huge variability among others in the literature, and in many instances, they are not fully reported. However, there are consistent patterns in the measures resulting from these dimensions. CAFE does not propose a differing set of dimensions and measures as many researchers in the domain tend to do, but rather enables the researcher to easily choose their own based on a gestalt overview of the domain.

Measures	Academic, Entertainment, Commercial Websites, User Generated Content	News Websites				
		Website	Source	Message	Medium	Sponsor
Credible	<u>TVE10**</u> , <u>CTD09†‡</u> , <u>AC09</u>	<u>CNS12#</u> , <u>CR02†‡</u> , <u>RH07*</u> , <u>S04 </u> , <u>BL11†</u> , <u>EMN06†‡</u> , <u>CTD09†‡</u> , <u>NC12†</u>	<u>FM03</u>		<u>CWL06*</u> , <u>MD08*</u> , <u>NC12†</u>	<u>FM07</u> , <u>JW09</u>
Expertise	<u>AAL13†</u> , <u>LWH14†</u> , <u>CTD09†‡</u>		<u>JF08</u>			
Trust, Trustworthy, Trustworthiness, Can be Trusted (G03††), Can be trusted/Cannot be trusted (WWPH08),	<u>MFZ03*</u> , <u>AAL13†</u> , <u>EMN06†</u> , <u>CTD09†‡</u> , <u>LWH14†</u>	<u>FM07</u> , <u>FM03</u> , <u>K01*</u> , <u>MFZ03*</u> , <u>AGSDC02*</u> , <u>G03††</u> , <u>JW09</u>	<u>FM03</u> , <u>JF08</u>	<u>FM07</u> , <u>FM03</u> , <u>C14</u>	<u>FM00*#</u> , <u>JW09</u> , <u>J05</u>	<u>FM07</u> , <u>JW09</u>
Bias/Fairness	<u>TVE10**</u>	<u>JW09</u>			<u>JW09</u> , <u>J05</u>	
Depth, In-Depth, Depth of Coverage, Depth or Completeness (JK02),	<u>BL11†</u> , <u>JK14</u>	<u>JK98</u> , <u>JK00*</u> , <u>JK02*</u> , <u>JK04**</u> , <u>JK10</u> , <u>CKK10</u> , <u>JK00*</u> , <u>B03*</u> , <u>K09**</u>	<u>JF08</u>		<u>JK08</u> , <u>KT09</u> , <u>JW09</u> , <u>J05</u>	
Goodwill						
Believability	<u>MFZ03*</u> , <u>BL11†</u> , <u>JK14</u> , <u>FM11</u>	<u>JK98</u> , <u>JK00*</u> , <u>JK02*</u> , <u>JK04**</u> , <u>JK10</u> , <u>FM07</u> , <u>S99§</u> , <u>FM03</u> , <u>MFZ03*</u> , <u>JK00**</u> , <u>B03*</u> , <u>CR02†‡</u> , <u>JK00**</u> , <u>C07*</u> , <u>AGSDC02*</u> , <u>JK14</u> , <u>CLP06</u> , <u>JW09</u>	<u>SN01</u> , <u>JF08</u>	<u>FM07</u> , <u>FM03</u>	<u>FM00*#</u> , <u>NC12†</u> , <u>KT09</u> , <u>JW09</u> , <u>JK08</u> , <u>J05</u>	
Bias, Unbiased (CKK10), Lack of Bias (JK02), Biased, Bias reverse coded, (FM03) and (RK09)	<u>MFZ03*</u>	<u>FM07</u> , <u>S99§</u> , <u>FM03</u> , <u>JK02*</u> , <u>CKK10</u> , <u>MFZ03*</u> , <u>AGSDC02*</u> , <u>G03††</u> , <u>JK98</u>	<u>SN01</u>	<u>FM07</u> , <u>FM03</u>	<u>FM00*#</u>	

Figure 3-3 Closeup of the repository shown in Figure 3-2 showing individual measures on the left and the domains they were used in at the top

Symbol Key

- * = Study was a medium comparison study. Other aspects of credibility of types of website/mediums were also included in the study
- † = Measure was formed of contributing components
- ‡ = Measure was a contributing component
- § = Measures were chosen as a result of factor analysis – there was other measures but the factor analysis showed that these made up credibility or one of its sub measures.
- | = Primary measure
- || = Classified as 'secondary components of credibility' which means it might have been weighted as less important than the primary component "Credibility"
- ||| = Classified as a tertiary measure of Credibility
- # = Study is a comparison of online news channels
- ** = Study included Blogs
- †† = Measured using a Semantic Differential Scale
- ‡‡ = Study assessed Structural Features

Figure 3-4 Symbol key for the codes in the repository

Reference Key

- AAL13† = (Ayeh et al., 2013), [PDF](#)
- AC09 = (Alsudani & Casey., 2009), [PDF](#)
- AM09 = (Armstrong & McAdams., 2009), [PDF](#)
- GSDC02* = (Abdulla, Garrison, Salwen, Driscoll, & Casey, 2002), [PDF](#)
- BRAH06 = (Bates, Romina, Ahmed, & Hopson, 2006), [PDF](#)
- BL11 = (Bae & Lee, 2011), [PDF](#)
- BS07* = (Banning & Sweetser, 2007), [PDF](#)
- CCB06 = (Bracken, 2006) [PDF](#)
- B03 = (Bucy, 2003) [PDF](#)
- C07* = (Cassidy, 2007) [PDF](#)
- CLP06 = (Choi, Lee & Pan., 2006) [PDF](#)
- CKK10 = (Chung et al., 2010) [PDF](#)
- CNS12 = (Chung et al., 2012) [PDF](#)
- CR02 = (Choi & Rifon, 2002) [PDF](#)
- CS10 = (Chesneya & SU, 2010) [PDF](#)
- CW06 = (Choi, Watt & Lynch, 2006) [PDF](#)
- C14 = (Clerwall, 2014) [PDF](#)
- CTD09 = (Cugelman, Thelwall & Dawes, 2014) [PDF](#)
- DB04 = (Dutta-Bergman, 2004) [PDF](#)
- E01 = (Eastin, 2001) [PDF](#)
- EMN06 = (Eastin, Yang & Nathanson) [PDF](#)
- FM00 = (Flanagin & Metzger, 2000) [PDF](#)
- FM03 = (Flanagin & Metzger, 2003) [PDF](#)
- FM07 = (Flanagin & Metzger, 2007) [PDF](#)
- FM11 = (Flanagin & Metzger, 2011) [PDF](#)
- FM13 = (Flanagin & Metzger, 2013) [PDF](#)
- FS04 = (Freeman & Spyridakis, 2004) [PDF](#)
- FS09 = (Freeman & Spyridakis, 2009) [PDF](#)
- G03†† = (Greer, 2003) [PDF](#)
- HS09 = (Hu & Sundar, 2009) [PDF](#)
- H06 = (Hong, 2006) [PDF](#)
- TH06 = (Hong (b), 2006) [PDF](#)
- JAZQ13 = (Jensen, Averbec, Zhang, & Wright, 2013) [PDF](#)
- J05 = (Jo, 2005) [PDF](#)
- JK19 = (Johnson & Kaye, 1998) [PDF](#)
- JK00 = (Johnson & Kaye, 2000) [PDF](#)
- JK02 = (Johnson & Kaye, 2002) [PDF](#)
- JK04 = (Johnson & Kaye, 2004) [PDF](#)
- JK09 = (Johnson & Kaye, 2009) [PDF](#)
- JK10 = (Johnson & Kaye, 2010) [PDF](#)
- JK07 = (Johnson & Kaye, 2007) [PDF](#)
- JK08 = (Johnson & Kaye, 2008) [PDF](#)
- JK14 = (Johnson & Kaye, 2014) [PDF](#)
- JF08 = (Johnson & Fahmy, 2008) [PDF](#)
- JW09 = (Johnson & Wiedenbeck, 2009), [PDF](#)
- K01 = (Kiousis, 2001) [PDF](#)
- K10 = (Kang, 2010) [PDF](#)
- KT09 = (Kim & Johnson, 2009) [PDF](#)
- KJ11 = (Kaye & Johnson, 2011) [PDF](#)
- KJ11(b) = (Kaye & Johnson, 2011) [PDF](#)
- KSM14 = (Kusumasondjaja Shanka & Marchegiani , 2014) [PDF](#)
- LC06 = (Long & Chiagouris, 2006) [PDF](#)
- LWH14 = (Lowry, Wilson & Haig, 2014) [PDF](#)
- SP08 = (Sweester & Porter, 2008) [PDF](#)
- MBP07 = (Mack, Blöse & Pan, 2007) [PDF](#)
- MD08 = (Melican & Dixon, 2008) [PDF](#)
- MC12 = (Morris et al., 2012) [PDF](#)
- MTH09 = (Menchen-Trevino & Hargittai, 2009) [PDF](#)
- MFZ03 = (Metzger, Flanagin, & Zwarun, 2003) [PDF](#)
- MR05 = (Moore & Rodgers, 2005) [PDF](#)
- NC12 = (Nah & Chung, 2012) [PDF](#)
- SO04 = (Okazaki) [PDF](#)
- RK09†† = (Rains & Karmikel, 2009) [PDF](#)
- RH07 = (Robins & Holmes, 2007) [PDF](#)
- RHS10 = (Robins, Holmes & Stansbury, 2010) [PDF](#)
- S99 = (Sundar, 1999) [PDF](#)
- SS08 = (Stavrositu & Sundar, 2009) [PDF](#)
- SN01 = (Sundar & Nass, 2001) [PDF](#)
- WWL04 = (Walther, Wang, & Loh, 2004) [PDF](#)
- WWP08 = (Wang et al., 2008) [PDF](#)
- W13 = (Westerwick , 2013) [PDF](#)
- WNB12 = (Willemsen, Neijens & Bronner, 2008) [PDF](#)
- XKL12 = (Xiea et al., 2012) [PDF](#)
- X14 = (Xu, 2014) [PDF](#)
- Y07 = (Yang, 2007) [PDF](#)

Figure 3-5 Key for the studies shown in the repository with a link to the PDF of each

3.6.3.3 *Inclusion in the Repository*

The repository does not include dimensions of credibility put forward as a result of factor analysis breakdowns of interviews or surveys conducted into how users perceive credibility or literature reviews of the domain, as many of these have not been used in an experiment to measure credibility (C. Gaziano & McGrath, 1986; Self, 2008). It only includes actual measures used in experiments. Users of CAFE are encouraged to consult with this repository when deciding which dimensions and measures to use when measuring the construct of credibility. The accompanying experiment platform also supports the use of multiple measures, in multiple configurations, and stores the resulting experiment data separately for detailed analysis. Thus, it is easier to analyse the impact of the intervention or undertake comparisons by individual dimensions and measures of credibility such as trust, expertise, accuracy, bias, fairness, depth of coverage etc. to derive additional findings.

3.6.3.4 *Using the Repository*

For each measure, the gestalt overview provided by the repository enables the researcher to identify: how often it was used, by which researcher, what other measures were used in the same study, which dimension of credibility, source, medium, or message was being measured, the domain the experiment was conducted, and how the final credibility score was derived. This can differ significantly depending on whether the researcher is attempting to measure or compare source, medium, or message, and the type of domain they are investigating e.g. news or health. By presenting the information in a table, users can quickly ascertain e.g. the most common measures on health websites, or how often, or in what domains e.g. 'Integrity' is commonly used. It also allows them to quickly see concentrations in activity.

This table forms the most comprehensive list of credibility measures in the literature and will serve as a reference point for future research. It will also help future researchers to adopt new and more appropriate measures, and perhaps help reduce the proliferation of often unnecessary additional measures which has blighted the domain.

3.6.3.5 *An Ongoing Classification Task*

As part of the development of CAFE an ongoing classification task has been instigated to record all the measures of credibility used in peer reviewed research. The repository website has been developed to store this information for easy access, and because it can be easily added to as time allows and new studies are published. As new studies are identified they will be added to this repository. However, at the moment, the search for new literature has been superseded by the need to redevelop the repository. Due to its growing size and complexity, it is now necessary to add a database backend to store the data. This will facilitate the storage of additional information and allow for additional functionality such as faceted search and visualisations. Once this redevelopment is complete, the latest group of studies will then be added to the repository..

3.6.4 VAS: An Improved Measurement Scale

Most studies on credibility utilise Likert or Likert-type scales to measure credibility or any of its individual dimensions and measures. The CAFE framework introduces the Visual Analogue Scale (VAS) which produced a more powerful form of parametric data and enables more fine-grained measurement.

Most of the research in the domain utilises Likert, or Likert type scales, Dials, Semantic Differential Scales, or Bipolar Adjective Rating Scales. For the most part, their use is a hangover from early paper-based data collection. As most research into credibility, and its individual dimensions and measures, is now undertaken digitally, a more appropriate measurement scale can be used. CAFE introduces a new a more powerful, yet fine-grained approach, the VAS, see Figure 3-6. To the best of our knowledge, no research into credibility has utilised VAS and they are introduced as a novel contribution to the domain.



Figure 3-6 Mono-directional Visual Analogue Scale (VAS) to detect the amount of Bias, 0 – 100, in a news article.

VAS were originally described by Hayes and Patterson and popularised by Aitken as a new technique for measuring pain, and as a communication aid between the patient and the clinician (Aitken, 1969; Hayes & Patterson, 1921). Originally, medical patients were presented with a paper-based ten-centimetre long VAS from 1 to 10 and asked to indicate their level of pain on it. Readings were taken by measuring in millimetres the distance of the indicated level of pain severity from either end of the scale, thus providing interval type data. Aitken turned to VAS due to a perceived concern in the accuracy of communicating a subjective personal feeling due to a scarcity of suitable terms in common speech to describe subjective experiences such as pain. The recording of a subjective hypothetical construct such as credibility, or any of its dimensions and measures, is analogous to the recording of a subjective personal feeling such as pain. VAS have also been shown to be reliable and valid instruments for the measurement of quantitative psychological tests of mood (Folstein & Luria, 1973). The CAFE framework and accompanying platform recommends and supports digital VAS, either in mono or bi-directional form, to record the user's perception of each of the measures chosen by the researcher to encapsulate the construct of credibility. The benefits include, a more powerful form of interval data, increased accuracy for small nuances, a wider range of scores, and less confusion with language or terminology. The result is continuous interval type data, a form of parametric data where the variance between the homogeneity of each of the tested designs

is considered the same. To improve validity and reliability, no metering, ticks or decorations should be applied to the scales (Matejka et al., 2016). VAS has recently and successfully been used in two implementations of the framework in two studies into bias, a core dimension and measure of credibility, see Figure 3-6.

3.6.5 *The Experiment Matrix, Latin Cubes, and Cuboids*

While Latin squares have long been used in experiment design, CAFE recommends Latin Cubes and Cuboids. This enables researchers to measure credibility and its individual dimensions and measures separately. By designing paths through each Latin cube or cuboid, the steps of which are then randomised, it is possible to ensure that each participant experiences each website/webpage, distortion, and measure once, thus achieving *incomplete counterbalancing*²⁷.

A core component of any experiment is the formation of the factors and levels that constitute all the combinations of the subject matter being tested. Latin squares were studied as early as the eighteenth century by Euler but were popularised in experiment design by Fisher, who introduced randomisation, and have been used extensively since then (Cox & Reid, 2000; Fisher, 1935). Latin squares are common in experiment designs and have regularly been used in credibility experiments. Latin cubes are used when the researcher needs to add another dimension to an experiment. While they significantly increase the complexity of a study, they also provide significant additional avenues for research. The authors know of no published credibility research which uses a Latin cube. As such, recommending their use forms a novel contribution to the domain. CAFE and the accompanying credibility experiment platform supports using a Latin cube in reduced from as the experiment matrix (Kishen, 1942). This enables the researcher to achieve incomplete counterbalancing by randomising the order of the steps in paths through the Latin cube. A Latin cube is a three-dimensional matrix of $N \times N \times N$ design. In CAFE, the matrix is typically made up of Websites/Webpage x Distortions x Measures combinations, henceforth known as WDM combinations, see Figure 3-7.

²⁷ Incomplete Counterbalancing is defined by the APA as: “an experimental design that controls for order effects by using a limited number of possible sequences of treatments administered in such a way that each treatment appears equally often in each position. For instance, the following arrangement of sequences of three treatments (A, B, C), each assigned to a different subgroup of participants, demonstrates incomplete counterbalancing: A-B-C to Subgroup 1, B-C-A to Subgroup 2, and C-A-B to Subgroup 3.” Source - American Psychological Association Dictionary: <https://dictionary.apa.org/incomplete-counterbalancing>

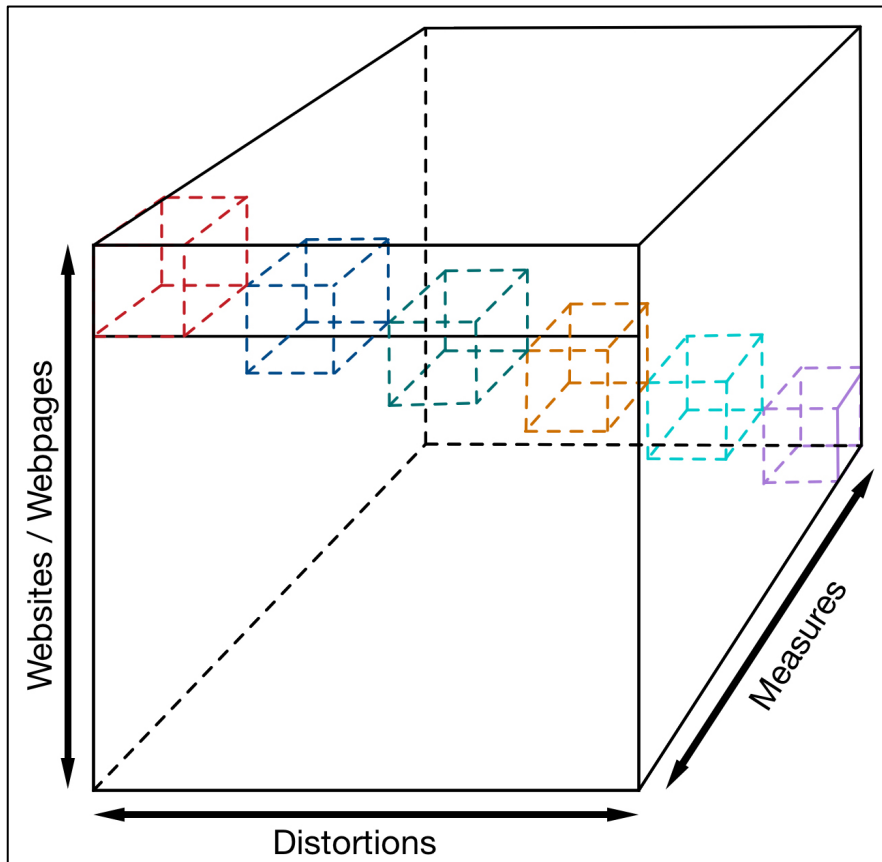


Figure 3-7 Isometric projection showing path 1 of 36 paths a participant might be assigned to in a 6x6x6 Latin Cube experiment. Each box represents a Website x Distortion x Measure (WDM) combination.

The following explains some key terms:

- *Website/Webpage*: This is the subject matter used in the experiment. This is an experiment factor, a controlled independent variable. One scenario might be an investigation into six different branding variations on six different news websites. Another scenario might involve six different types of webpages from the same website.
- *Distortion*: These are the variations of the websites/webpages being tested and form the second experiment factor. The distortion process may also be known as the treatment. The resulting distorted versions of each websites/webpages then form levels in the experiment. In the above first scenario, the six distorted versions of each website would have different logos and branding applied to the same website design to see if any impacted perceived credibility.
- *Measure*: The choice of and quantity of measures used is dependent on the dimension of credibility, source, medium, or message that is being measured and the domain it is being measured in, e.g. news or ecommerce. It also depends on the definition of credibility the researcher uses to encapsulate the construct and that they want to convey to the participant. As highlighted earlier, researchers have used anywhere between one and twenty-one measures to create an overall credibility score. These are the dependent variables in the experiment.

- *WDM Combination:* A Website/Webpage x Distortion x Measures or WDM combination is the result of the intersection of the three independent and dependent measures in the Latin cube or cuboid. The result forms a single credibility rating task in the experiment. Each participant is shown a series of these based on the path they were assigned to in the Latin cube or cuboid. The website/webpage may be live, but is usually displayed as a static image. One or more distortions are applied to each website/webpage so that there are different versions of each. This allows the researchers to investigate or compare the impact of different technical or structural features or characteristics of the overall design, such as different layouts, colour schemes etc. An original or undistorted version of a website/webpage may also be included as a control. Depending on how the experiment is setup, researchers may decide to use one or multiple measures. If they choose to use multiple measures, they may use them independently, thus adding the third axis which can be seen in Figure 3-7. Alternatively, they may ask the participant to rate the credibility of a webpage/distortion combination with multiple measures at the same time, thus the third axis will be one deep.

For more complex experiment designs necessitating four or even five factors, superimposed cubes add additional orthogonal dimensions and are known as Graeco-Latin cubes and Hyper-Graeco-Latin cubes respectively. While the individual cubes are of first order, their combination is of second order. Latin cubes are best suited to situations where a factor has multiple levels, in this case the resulting distorted website/webpage, and the researcher believes that there are little or no interactions between factors, in this instance websites, than a full factorial design is (Hill & Lewicki, 2006).

3.6.5.1 *Balance in the Experiment Design*

The major benefit of an experiment design using equal amounts of websites, distortions, and measures is that researchers can design every path through the resulting Latin cube experiment matrix so that each participant experiences each website, distortion, or measure just once, see example path in Figure 3-7. While a balanced Latin cube, is ideal, due to research requirements, it is not always possible. Consequently, a Latin cuboid, can also be used. This enables a researcher to add additional websites/webpages, distortions, or measures as required. When a Latin cuboid is employed, researchers should be cognisant that participants may see one or more of the websites/webpages, distortions, or measures more than once, or not see some of them at all, depending on how the path through the matrix they are assigned to is designed.

Alternatively, researchers may decide to forgo having multiple independent measures which the participant sees individually. Traditionally most researchers in the domain have presented participants with all the measures at the same time. This was primarily due to a hangover from older paper-based elicitation techniques, and to avoid the huge increase in time, complexity and overheads because of the additional participants required if they were presented separately. Recently, most research has been undertaken online

which enables more complex experiment setups to be more easily devised and deployed. The introduction of crowdsourcing and the ease of access and relatively low cost of participants has also made it much easier to present the measures individually. Showing participants all the measures at the same time may also have some drawbacks. First, displaying too many measures for each WDM combination may induce fatigue in the participant. As stated earlier, researchers have used anything between one and twenty-one measures for the construct of credibility. Second, participants may not be able to differentiate between individual measures such as fairness and bias which may increase frustration. Third, to ensure rigor, the order of measures should be randomised to ensure that the measures displayed at the top are not favoured when large amounts are employed, adding complexity to the design of the experiment platform. Lastly, as maintained by Eastin noting previous work by Gaziano and McGrath, the choice of measures may influence perception and understanding of the construct (Eastin, 2006; C. Gaziano & McGrath, 1986). By presenting the measures separately, the focus is on the measure, e.g. trust, rather than the arch-concept, credibility. These independently gathered scores could then be combined as they would if they were gathered at the same time. It should be noted however that this method requires further validation and reliability testing.

Conversely, showing participants measures individually provides a singular benefit. Trust, accuracy, expertise etc. are all important individually and as core dimensions of credibility in domains such as health, government, ecommerce etc. Consequently, having pure or unadulterated scores for each measure would be score beneficial to website designers or researchers. It also makes it easier to calculate a weighted credibility depending on the importance of the measure and the domain. This is further detailed in section 3.6.6.

3.6.5.2 *Paths Through a Latin Cube Experiment Matrix*

By defining paths through a Latin cube, it is possible to ensure that each participant sees each WDM combination once. This ensures that participants do not get familiar with any websites/webpages or distortions in the experiment. Because a Latin cuboid is unbalanced, the path can be defined so participants either see each WDM once but do not see every WDM, or they see some WDM more than once.

To devise a path, the first step of each path should be selected from a different one of the intersecting WDM combinations in the top layer of the Latin cube. Figure 3-7 depicts a single path through a 6x6x6 Latin cube. Figure 3-8 is an alternate representation of the same Latin cube showing 6 of the possible 36 paths through it. The same Path 1 is shown in Figure 3-7 and in Figure 3-8. In each representation, step one is the bottom left WDM combination of the top level of the Latin cube. The next step in each path should be plus one on each axis of the NxNxN cube. E.g. If step one, of path one, is NxNxN, step two, of path one, is (N+1)x(N+1)x(N+1). Essentially, each path cuts a diagonal line, on three dimensions, through the Latin cube. When the end of an axis is reached, it loops back to the start of that dimension. This can be seen in Figure 3-7 between step 5 and step 6 in path 2 where the dimension's axis (D) loops back from 6 to 1,

W5D6M5 – W6D1M6. Every WDM combination in the Latin cube can thus be assigned to a single path. The number of paths necessary to intersect with every WDM combination in a balanced Latin cube is derived by multiplying two axes of the cube. E.g. The 6x6x6 Latin cube represented in Figure 3-7 and Figure 3-8 requires 36 paths (6x6), each 6 steps long, to ensure equal and full coverage of each WDM combination. If a Latin cuboid is being used, then the researcher has two options to make sure the additional WDM combinations are covered in a balanced manner. 1) They can decide to extend the paths by looping them back around to cover the additional WDM combinations. The main issue with this approach is that participants will see one or more of the websites, distortions, or measures more than once. Though unlikely, the carryover effects of this experience may impact subsequent credibility rating tasks in the same experiment. 2) They can devise a range of shorter paths so that each participant does not see every WDM combination, but the ones they do see are never repeated. The shorter paths make it easier to ensure that each WDM combination is covered.

Once all paths have been devised, participants should be assigned to them at random. Many statistical methods, such as Two-Way Repeated Measures ANOVAs, require balanced or roughly balanced amounts of data. Therefore, it is important to ensure roughly equal numbers of participants are assigned to each path. This can be achieved by dividing the total number of users derived from the Statistical Power Analysis test with the total number of paths and using this as a cap on each path. When the number of participants assigned to the path meets this cap, the path is removed from contention. Once a participant is assigned to a path, the order of the steps in it should also be randomised, thus achieving incomplete counterbalancing. This reduces the impact carryover and fatigue effects on the experiment. The experiment platform accompanying the CAFE framework supports all these existing best practices, recommendations and features. By adding a control WDM combination, to each path in the experiment matrix, the experiment takes the form of a series of multiple single variant A/B tests where A is the control and B is the range of distortions (Siroker & Koomen, 2013).

Another approach to ensure validity and to add an additional avenue of exploration to an experiment is comparative re-evaluation. By recording the WDM combinations that participants experience, and then displaying the same combinations to them for comparative re-evaluation, participants are given another opportunity to confirm their ratings. This approach has proven successful in two implementations of the CAFE framework and platform to measure the impact of structural features and wider characteristics of the design on the perception of bias in news websites (Spillane et al., 2017a, 2018). This approach adds an additional orthogonal dimension of user ratings to the Latin cube which can then be compared to the original. A major benefit of this is the opportunity to discover which websites, distortions, or measures are more likely to be amended and by how much. It also helps allay a key concern, namely rapid fire clicking, during crowdsourced experiments. The main supposition behind this is that participant's may want to revise their initial ratings on measures such as trust, bias or authority for webpages with a particular visual aesthetic when they see them alongside other webpages. The impact of a single news website design on any of these measures may be nominal until multiple types of webpages and designs are compared, such as those of tabloid and broadsheet newspapers.

Another concern researchers should be aware of is confounding variables. The most obvious manifestations of this in online credibility research would be, previous experience or knowledge of the topic, the organisation or brand behind the website, and personal biases. As much as possible, the experiment design should consider or attempt to negate the impact of confounding variable or take them into account during statistical analysis as covariates.

3.6.5.4 Accounting for the Impact of Textual or Visual Content on Experiments Focusing on the Impact Website Design or Visual Presentation on Perceived Credibility

A major concern that researchers should be cognisant of is the impact of textual or visual content in a website or webpage in studies comparing or investigating individual technical or structural features, or the overall characteristics of a website's design or visual presentation. If researchers use dummy content such as the lorem ipsum, then the experiment will not reflect realistic browsing. Alternatively, if the same content is used in each webpage, then participants will get bored and become less engaged as the experiment proceeds.

One partial solution is to assign different individual pieces of content, news articles etc., to individual webpages for the duration of the experiment. Alternatively, they could be assigned to each distortion for the duration of the experiment. By assigning participants to paths through the Latin cube or cuboid that ensure that they experience each website, distortion or measure once, they will always experience new content in each credibility rating task, thus more closely replicating a real browsing experience. However, in both options, the content may act as a confounding variable, thus reducing the effectiveness of the experiment.

Consequently, to complete the above partial solution, the researcher may use different individual pieces of content of a common, or varying credibility. Either way they must account for its impact. This can be achieved in a pre-test with a small audience, but one which should have a similar profile to the intended experiment audience. A process to create and pre-test a range of biased news articles was successfully used in one of the deployments of CAFE and the accompanying experiment platform (Spillane et al., 2018). Thus, it is possible to use the pre-test content credibility scores as covariates in the final statistical analysis.

It is also worth considering replacing one of the WDM combinations with a plain text version of each piece of textual content, without any website design or aesthetic treatment in the experiment. This will establish a ground truth credibility rating for each piece of content, which may also have to be considered during subsequent statistical analysis.

3.6.6 A New Method to Derive a Credibility Score

Another novel contribution of CAFE is a proposal for a new method to calculate credibility scores by weighting each of its dimensions and measures based on their importance to the construct of credibility as it is encapsulated by the researchers and participants, and depending on the domain it is being measured in. This proposal is detailed in section 9.11 of the appendix.

3.6.7 Statistical Testing

Choosing an appropriate statistical test depends on the experiment aims, what the resulting data supports, and the requirements for reporting. It is important that researchers consider these during the early experiment design stages. Considerations include the number of factors in the experiment, or the choice of measurement scale etc., which all impact the type of statistical analysis that can be undertaken. CAFE is designed to make researchers plan the statistical analysis early in the experiment process. Further details on this can be seen in section 9.12 of the appendix.

3.6.8 Increased Reporting of Individual Measures and their Impact on the Credibility Score using Visual Aids

The final best practice recommendation and novel contribution of CAFE focuses on how credibility scores are reported. Based on a detailed review of studies in the domain, several common failings were noticed. These include not reporting the Mean, Standard Deviation, and Variance of individual measures. Further details with links to example studies are provided in section 9.13 of the appendix. CAFE makes recommendations to address these and to suggest using box plots or scatter graphs to display the results of individual measures of credibility separately so that the impact of each dimension and measure can be derived. These are also further detailed in section 9.13 of the appendix.

3.6.9 Target Audience and Use of the CAFE Framework, Experiment Platform, and Recommendations

The eight existing best practice recommendations and novel contribution highlighted thus far are designed to improve different stages of a typical credibility experiment, see Figure 3-1. They arose from an in-depth review of the literature, and in many cases, grew out of frustration from reading or trying to compare different studies.

The target audience for this framework is novice researchers, or those new to the domain, to help them identify issues with their experiment design early in the design process so that they can be addressed before they inhibit or invalidate an experiment. The accompanying experiment platform will also help technical researchers (i.e. those who are familiar with computer programming, software environments, and the ability to deploy code to servers) and non-technical researchers (i.e. those without, or with limited ability to program, and without the skills to connect to servers to deploy code) to quickly deploy experiments.

3.7 Design, Development, and Implementations of the CAFE Framework and Experiment Platform

The CAFE framework was developed in tandem with the experiments conducted as part of this thesis, and it will continue to be enhanced afterward. Each experiment was an opportunity to evaluate and modify the framework to ensure that it was generic enough to support the design and development of multiple types of experiments on credibility and its individual dimensions and measures such as bias. The framework and platform have also recently been successfully utilised and implemented in final year undergraduate and Master's level research dissertation projects.

3.8 The CAFE Framework and Platform: Evaluating Case Study Implementations with Different Experiment Configurations

The CAFE framework and accompanying experiment platform was created for the purposes of conducting the three experiments as part of this thesis, thus its evaluation is its three successful deployments for each of the three experiments. It has also been evaluated during its use by students in their research projects and continues to be used for the design, development, and deployment of experiments on bias and credibility outside the scope of this thesis.

As this is the first empirical framework in the domain to support researchers in the design, development, and deployment of experiments investigating credibility online, or any of its individual dimensions and measures, there are no alternatives to compare it to. Consequently, the most appropriate evaluation of

CAFE is its recent successful implementations and continuing usage. However, three facts must be acknowledged.

- It has mainly been used by its creator or by students conducting research projects guided and supervised by its creator. Their thorough understanding of CAFE may make it easier for them to deploy compared to future users.
- Despite their best efforts, there is always the possibility that its design has been over fit to the author's needs and may not immediately match the needs of others. As a result, the framework may need to be extended and refined, based on the evaluation of other researchers in the domain
- Thus far, only five studies on bias, a core dimension and measure of credibility, have been completed using the CAFE framework and experiment platform. A sixth study on credibility is near to completion. However, CAFE is also being used successfully to aid in the design of several further experiments on credibility in multiple domains, some of which are close to deployment.

Despite these, the first version of the CAFE framework and accompanying platform have performed well in recent deployments. Future versions will be expanded with additional features and more documentation.

In future, a formal evaluation will be undertaken of its use by other researchers, who will be tasked with using it to design and deploy experiments on credibility, or any of its dimensions and measures such as bias, trust, expertise etc. This formal evaluation will likely consist of evaluations of satisfaction, usefulness, ease of use, utility, support, etc. which are all commonly used as measures of such systems and processes. via structured or semi structured interviews. However, this cannot occur until the framework and the accompanying experiment platform are made public and its use promoted. This is also out of scope for this thesis. For the purpose of this thesis, the evaluation of the framework and accompanying experiment platform is its successful use and deployment for the three experiments conducted as part of this thesis, the two experiments completed by two Masters and one 4th year student.

3.9 Discussion

The CAFE framework and accompanying experiment platform are designed to simplify and formalise the design, development, and deployment of experiments investigating the impact of website design or webpage aesthetics on perceived credibility, or any of its dimensions and measures. It was developed and refined through multiple studies on bias, a core dimension (Cockburn et al., 2018) and measure of credibility. It has also been further developed and refined during the design and development of several ongoing studies on credibility.

CAFE is an open framework and researchers are encouraged to adapt it to suit their individual experiment aims, to better consider how each decision they make affects the later stages of the experiment lifecycle. It should prompt discussion and debate over different strategies and their merits. It is designed to make sure

a team of researchers are all in agreement and have the same understanding of all aspects of an experiment. It also acts as a record of the experiment design decisions taken by the team prior to the experiment being run. Thus, potential issues can be discussed and ironed out at an early stage rather than when it is too late, and researchers are forced into an alternative, less effective strategy. The design decisions, which are recorded by the researcher in their implementation of the framework, as described in section 3.3.2, can be made available as part of the preregistration of experiment plans and hypothesis on platforms such as OSF²⁸ 3.3.2, and subsequently as part of the FAIR data publishing campaign to improve repeatability in scientific experiments (M. D. Wilkinson et al., 2016).

A major component of CAFE is the online repository of measures. This taxonomy identifies the dimension of credibility, source, medium, or measure, being investigated in each study, it also identifies the type of study, the type of website, and how the credibility score was derived. This is the first of its kind and involved an extensive trawl through the literature. This is also an ongoing effort. As new studies are discovered they will be added to this repository. The identification of appropriate measures of credibility has been a major issue in the past. The sheer scale and the volume of literature in the domain means that it is virtually impossible for researchers to read all the existing work while also attempting to stay up to date with new work. Consequently, researchers have been guilty of duplicating or defining new dimensions and measures, when there is little evidence the existing measures would not have sufficed. To be clear, there are many fine examples of excellent research identifying dimensions and measures (Cugelman et al., 2009; Eisend, 2006; McGrath & Gaziano, 1986), constructing scales (Appelman & Sundar, 2016), and undertaking rigorous scientific investigations into credibility in specific domains and contexts (Chung et al., 2010; Flanagin & Metzger, 2007; Hu & Sundar, 2009; Sundar & Nass, 2001). This research recognises these and other significant contributions but argues that a central resource and repository would make identifying such easier. A significant benefit of the repository is the gestalt overview of the domain afforded by categorisation. This classification and repository are being expanded to include additional information about each study, such as the number of participants, and to add additional functionality such as faceted search and visualisations.

This repository is not designed to stifle or limit which measures should be used. Since the work of Hovland and his team in the 1940s, credibility has typically been defined as a multi-dimensional construct, the dimensions of which change depending on a multitude of factors (Hovland, 1948; Hovland et al., 1949). Many researchers go to great lengths to define accurate dimensions and use one or more measures and scales to quantify them. The dimensions chosen by the researcher also conveys their encapsulation of the construct to the participant. However, not all researchers go to such great lengths, and many simply choose common measures from the huge range within the literature, without necessarily thinking about the nature

²⁸ <https://osf.io/>

of the construct conveyed to the participant and whether the participant's encapsulation of the construct matches what the researcher is trying to measure.

The aim of the repository is not to make it easier for researchers to simply choose the most common measures, or to encourage them to choose the same measures that have been used in similar studies. The aim is to provide them with a gestalt overview of the range of measures, so that they might more keenly consider which ones they should use. The researcher should then follow the relevant links to the associated research papers and read about how and why each dimension and measure was used. Of course, they may end up using the most common measures, or the same as other studies, but their priority should be to use the most appropriate. This classification and repository help to solve the issue of the unknown unknown highlighted earlier, where researchers are only aware of the measures contained in the studies they have read and therefore may not be using the most appropriate for their purpose.

The repository might instigate deeper thinking by researchers into empirical investigations into credibility. Firstly, it may instigate validation and reliability checks on many of the measures and result in a clean-up of the domain. Secondly, as shown in section 2.2.4, there are four types of definitions of credibility, perceptual variable, believability, dimensions, and philosophical. Yet the vast majority of studies only attempt to measure it as a product of its dimensions. The variations in the methods of measuring credibility apparent in the repository may make some researchers to consider alternative means altogether to measuring the construct. This may include measurement techniques based on the other three types of definitions. Alternatively they may take inspiration from the few existing examples of alternative approaches which were briefly highlighted in Footnote 9 on page 50.

The repository may also instigate deeper thinking on new and alternative methods of measuring bias. Firstly, as has been previously shown, much of the empirical work on bias focuses on factors surrounding the text of a news article, the spoken word of radio, or the visual of television, or the message contained in them. It must be acknowledged that this is a failure of researchers in the domain, and new methods need to be developed for measuring bias within the text, spoken word and/or visuals. Inspiration for this may be found in some of the credibility studies in the repository which focus on content. Secondly, researchers may be inspired to investigate and identify the dimensions of bias, whether this is the dimension of bias being investigated, or the dimensions of the construct itself. Thirdly, one of the biggest challenges in the domain is the lack of clear and obvious methods to deal with researchers and participants own biases.

However, it remains to be seen whether it acts as a catalyst to encourage such deeper thinking on the subject and as a reference point for further reading in the domain, or as a shortcut enabling researchers to simply pick the most common approaches used and/or those in the most similar domain.

Another point of discussion is the accompanying experiment platform. This has been released to the community to aid future researchers, with and without a technical background, to quickly deploy experiments online measuring, comparing, or investigating credibility or any of its individual dimensions and measures. It is designed to be used with the CAFE framework and supports all the aforementioned experiment recommendations. It must also be acknowledged that such a platform may actually limit, or hamstring innovation, as researchers may be encouraged to design experiments to fit its capabilities rather than what they want to achieve. Nevertheless, it is hoped that it will prove a boon to many.

The CAFE framework, the accompanying platform, the repository and classification of credibility measures, and the experiment design best practices and recommendations, and the other research contributions are also put forth with three acknowledged facts. First, there are many researchers in the domain vastly more experienced who already practice high standards in experiment design and research. To these, the benefits of the CAFE framework are limited when compared to novice researchers. There are however significant other contributions which will likely prove useful. Second, the experiment design best practices, recommendations, and novel contributions in no-way claim to be the only solutions, they are simply put forth to help solve or contribute to solving existing issues. Third, this is best suited to novice or new researchers to the domain interested in research into the impact of website design or visual aesthetics on perceived credibility or any of its dimensions and measures such as trustworthiness, expertise, fairness, accuracy, bias, depth of coverage etc.

3.10 Conclusion

CAFE is an open framework designed to help researchers design, develop, and deploy experiments investigating credibility online or any of its dimensions and measures. This has far reaching benefits for new researchers in the domain and will greatly aid reproducibility and longitudinal comparisons. It is designed to be adapted by researchers to map out their experiments to identify potential issues or conflicts early in the design process. The code-base for the accompanying experiment platform has also been released and it is intended that its capabilities are significantly expanded in future. It will greatly simplify the process of deploying experiments for those without a technical background or those new to the domain. The online repository of credibility measures is also a tool to assist future researchers in the domain. It is designed as a single reference point to simplify the process and reduce the workload for researchers to choose the most appropriate credibility measures for their study.

“It has been said that politics is the second oldest profession. I have learned that it bears a striking resemblance to the first.” Ronald Regan 1977

4 Experiment One – Perception of Bias: The Impact of User Characteristics, Types of News Website Design, and Individual Technical or Structural Features

4.1 Introduction

To investigate the impact of website design and website aesthetics on the perception of bias in news websites, it was decided to first look at specific features such as advertising, promoted content, and social media integration. Modern news websites are highly complex pieces of software fully integrated with the news agencies publishing platforms. News websites sit at the end of long publishing pipelines²⁹, the stages of which include; writing, fact checking, source confirmation, graphics and artwork, and editorial signoff, before finally going live. For the most part, the individual webpages within a news website have a common, visual appearance and structure derived from the main template which in turn is derived from the main design. Whether employing a responsive or non-responsive format, almost all the major news agencies have adopted a grid layout in their main template, whereby webpages are constructed from a series of ‘blocks’ some of which can extend across multiple columns in the grid. For the most part, each block is aligned to the top of the page, or to the bottom of the block immediately above it. Even if a news agency is using a responsive website design, which more easily accommodates news consumption across different browser set ups and devices, the design and results of this experiment are valid. In a responsive design, the variability in the number and width of columns does not negate the fact that each block is aligned to the blocks above it.

Apart from news articles, news webpages also contain a variety of standard technical features which are visually adapted by the website’s Cascading Style Sheets (CSS) to match the overall design. These technical features provide an array of services or information to the user. They include, search facilities, advertising, navigation, recommended content, article metadata, author profiles, related content, or interaction facilities such as social network sharing and comment facilities. Many of these features are contained within a block in the overall grid of each webpage, styled to fit in with the overall design.

²⁹ The Washington Post wrote an article about their new: “central nervous system” or “brains”, a “Mount Everest-sized bundle of computer code that controls virtually every word, photo, video and story”. Costing \$7 million in 2011 and taking 18 months to implement, the Content Management System (CMS) controls every aspect of their content delivery https://www.washingtonpost.com/opinions/behind-the-posts-redesigned-web-site/2011/03/25/AFC3GXYB_story.html?utm_term=.caa0f403ffa8

In this way, news webpages are constructed from a series of blocks containing standard technical features. Each news webpage contains some fundamental features such as the main banner, main and secondary navigation, main content area, footer etc. Then depending on a number of constantly changing variables and outside influences, the rest of the webpage is automatically constructed. These variables include; the section of the news website being explored, the content or topic of the news article, the user model - especially if the user has logged into the news website or they have integrated their social media accounts, trending news content, personalisation features, and cross website browsing and advertising tracking. Thus, when a user visits a news webpage, they may experience a completely unique combination of individual technical features, many of which, such as advertising, or promoted internal and external content, can display an almost unending range and combination of different material. Thus, each visitor to a news webpage may have a tailored or even unique experience.

As the individual technical features play such an important role in the overall visual presentation of a news website, it was decided to investigate whether they may impact the perception of bias. The impact of user characteristics such as age, sex, and political leaning would also be investigated at the same time as there is currently no literature on this. Lastly it was also decided to investigate whether different categories of news website, traditional print, news magazine, and international multi-format news agency, impact how users rate perceived bias. It was decided not to investigate tabloids as it was felt that including this genre would too heavily influence the perceived bias ratings and distort any potential findings.

4.2 Influence from the State of the Art

The main influences on this experiment are drawn from the requirements of the State of the Art identified in 2.2.12 and from the discussion section of the CAFE framework presented in section 3.9.

- 1) There is a wide range of research demonstrating that users' judgements of the credibility of information online is impacted by its visual presentation
- 2) A large body of theory exists in the domain explaining how users form judgements of credibility online, most of which highlight the importance of heuristics and the visual presentation
- 3) This is underpinned by long standing theory such as the ELM, HSM and the C&AHIP
- 4) The visual presentation has also been shown to impact trust, another core dimension and measure of credibility
- 5) Bias, like trust, expertise, accuracy, fairness etc., is a core dimension and measure of credibility
- 6) Its importance as a dimension and measure increases when users are judging news online, as opposed to e.g. health related, financial, or historical information
- 7) It is the contention of this thesis that users' perception of bias in a news article, like credibility, is affected by its presentation. One of the main underlying suppositions behind this research, is that due to the constant prevalence and availability of news on the WWW, users have adopted

peripheral, heuristic, or automatic methods of evaluating the news that they consume online via visual cues in its presentation

- 8) Currently, there is no existing research investigating the impact of a news website's design, reflected in the aesthetic of a news webpage, on the perception of bias in the news articles they contain. This is despite an extensive literature review, and the importance of bias to the producers and consumers of news. It is also an obvious gap in the body of knowledge
- 9) The findings from this research will be important to news consumers, news website designers, and online editors and journalists. It may also lead to the development of formal models explaining how users form judgements of bias online and in other domains
- 10) No theoretical models or empirical frameworks were discovered to aid in the design, development, and deployment of experiments to investigate the impact of the visual presentation on the perception of bias. Demonstrating the lack of research into this construct, a review of the aligned domain of trust discovered ten theories and five frameworks explaining how judgements of trust are made, and six empirical frameworks describing how it should be measured
- 11) A review of the overarching domain of credibility discovered ten human-centric models, theories, frameworks, and schematics explaining how judgements of credibility are made online, but no empirical framework was discovered
- 12) The creation of an empirical framework and experiment platform for the design, development, and deployment of experiments to measure credibility, or any of its individual measures such as bias, will have more utility in the domain

There were two specific reasons it was decided to investigate the impact of technical or structural features of a website's design or a webpage's aesthetic on the perception of bias. First, due to a range of similar work in the aligned domain of trust and some work in the overarching domain of credibility. For trust, this related work includes work on the impact of pictures (Roy et al., 2001), logos (P. B. Lowry et al., 2014), and colours (J. Kim & Moon, 1998). For credibility they include similar studies on social feedback (Giudice, 2010), seals of approval (Burkell, 2004), and inline or embedded hyperlinks (Borah, 2014). These are extensively detailed in sections 9.8.2 and 2.2.11 of this thesis respectively.

Second, many of MTFS reviewed earlier, highlight the importance that individual features of a website's design, which are reflected in a webpage's aesthetic, have on judgements of credibility. Bias is a core dimension and measure of credibility, especially when judging news online (Fogg et al., 2003). The MTFS are detailed in section 2.2.8, and the means by which they claim users may rely on individual features and characteristics are detailed in section 0. Figure 4-1 is a screen shot from the news website Independent.co.uk. Highlighted are different categories of surface level features by which readers may use as visual cues to make heuristic judgements of bias in the news article.

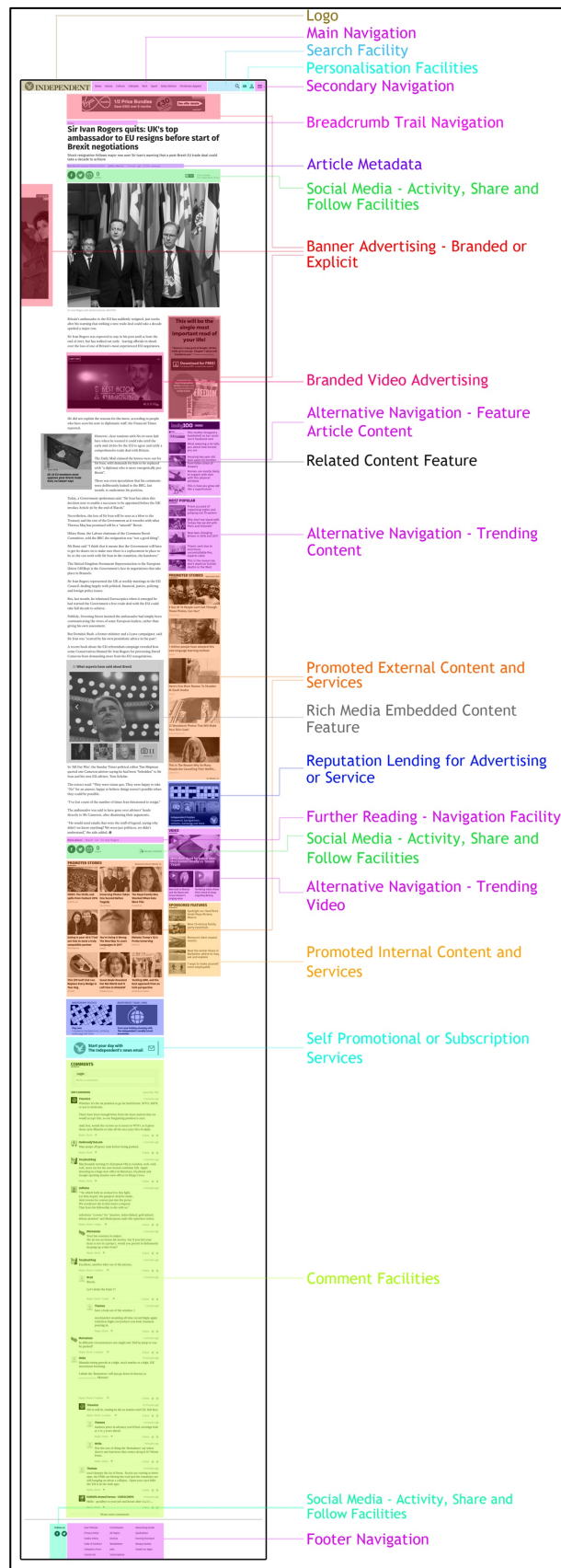


Figure 4-1 A webpage from Independent.co.uk with the main categories of technical features highlighted in colour.

4.3 Experiment Objectives and Contributions

4.3.1 Objectives

This experiment partially satisfies two of the research objectives. First, it provided an opportunity to test an initial version of the CAFE framework and the first iteration of the experiment platform as stipulated by:

- **Objective 2:** To design, develop, and test an innovative framework and platform that supports a range of experimental designs to detect and record user's perception of bias.

Second, the experiment was designed to satisfy the first clause of the third research objective, namely whether features of a news website's design, reflected in a webpage's aesthetic, impact perceived bias:

- **Objective 3:** To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impact perceived bias.

4.3.2 Contribution

This experiment provides evidence for the first half of the Major Contribution, namely the identification of features and characteristics.

- **Major Contribution:** The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias

4.4 Research Ethics Approval

Research ethics approval for this experiment was granted by the Research Ethics Committee of the School of Computer Science and Statistics, Trinity College Dublin.

4.5 Use of Website Titles to Describe Results and Findings

It should be noted that in this and the two other experiments described in this thesis, website titles such as The Guardian or Al Jazeera are commonly used to describe results, e.g. the BBC was the most or least biased. This was done for simplicity and to differentiate the nine distorted webpage/article combinations. Any results and findings are not reflective of the publication in reality, i.e. if a website such as Reuters is discovered to be the most or least biased in one of these experiments, that is most likely the result of the content inserted into each webpage and the distortions applied to it, though there is a chance some of it can be ascribed to participants recognising the news agency and their positive or negative feelings towards it.

4.6 *Experiment Design*

The experiment was designed using an early version of the CAFE empirical framework detailed in chapter 3. The process of designing this experiment contributed significantly to subsequent iterations of the framework. The experiment was set-up under an exploratory primary hypothesis to determine whether perceived bias in news articles could be impacted by distorting the visual presentation of news webpages by removing common individual features. The two independent measures were the webpage/article combinations and the distortions. The dependent measure was perceived bias. A 9x9 incomplete counterbalanced measures design was employed to test a broad range of webpage/article combinations and distortions. As each participant experienced multiple articles, broadly relating to the same topic, it was decided to test whether or not the subject of the article was portrayed with a positive or negative bias. Bias has been conceptualised in a positive v negative way by a range of researchers (H. S. Friedman et al., 1980; D. T. Lowry & Shidler, 1998; Miller et al., 2007; Moriarty & Popovich, 1991; Rada, 1996; Rainville & McCormick, 1977; M. J. Robinson & Sheehan, 1983). Following the main experiment, in two stages, participants would be presented with the 9 distorted webpage/article combinations they had just rated and asked to pick the most positive and the most negative. Participants would then be asked a series of quantitative and qualitative reflective questions.

4.6.1 *Experiment Hypothesis*

4.6.1.1 *Primary Hypothesis*

The primary research hypothesis the experiment was conducted under was designed in accordance with the exploratory objective.

H₀ Perceived bias in the news articles will not be affected due to distortions in the news webpage.

H_A Perceived bias in the news articles will be affected due to distortions in the news webpage.

4.6.1.2 *Secondary Hypotheses*

Three secondary hypotheses were constructed to focus on specific aspects; individual features, the category of design, and user characteristics.

4.6.1.2.1 *Individual Features*

H₀ Perceived bias in the news articles will not be affected due to removing individual features of the webpage.

H_A Perceived bias in the news articles will be affected due to removing individual features of the webpage.

4.6.1.2.2 *User Characteristics*

H₀ User characteristics will not affect perceived bias in the news articles.

H_A User characteristics will affect perceived bias in the news articles.

4.6.1.2.3 *Category of News Website Design*

H₀ Perceived bias in the news articles will not be affected due to the category of news website.

H_A Perceived bias in the news articles will be affected due to the category of news website.

4.7 *Methodology*

The experiment required participants to rate their perception of positive or negative bias in the news articles contained in nine news webpages. Each webpage/article combination was subject to one of eight distortions which removed individual features from the visual presentation. Along with a control, a 9x9 experiment was created. 135 participants submissions were included in the final experiment dataset from which the findings were derived.

4.7.1 *Experiment Content*

4.7.1.1 *Websites and Webpages*

A single webpage from nine internationally popular news websites were used in the experiment³⁰. The selection of these websites was based on the following criteria:

- 1) To represent the three main domains, traditional print, news magazines, and international multi-format news agencies. Tabloid websites were not used due to likely strong feelings about the category influencing results. This is backed up by the finding of the reflective questions in experiment one.
- 2) All the websites were English language. This was important as participants' primary language would be English and much of the text on each website is contained in graphics and images which would be difficult to edit/translate.
- 3) They were all British-based publications, except for Al Jazeera, which has a major London office. Al Jazeera was included as the third international multi-format news agency as no other British based news organisation fit that description, because of their large British presence, and finally because of the similarity of its website to the websites of the two other international multi-format

³⁰www.theguardian.com, www.telegraph.co.uk, www.independent.co.uk, www.economist.com, www.spectator.co.uk, www.newstatesman.com, www.aljazeera.com, www.bbc.com, and www.reuters.com

news agencies in the study, the BBC and Reuters. As participants were to be pre-screened to make sure they were from the US, news websites from the US were not used to help negate any strong feelings about individual news agencies, such as the liberal MSNBC and Huffington Post, or the conservative FOX News and Washington Times, in case they were recognised. While it is still possible that the US participants could have strong feelings towards UK based news websites it was felt that it was less likely.

- 4) They are among the most popular news websites within their respective categories, with the supposition that the most popular news websites would be good examples of professional design and aesthetics.
- 5) They belonged to quality press³¹ or a serious news agency rather than content farms or tabloids.
- 6) Their design was sufficiently complex, and for the most part had large amounts of advertising and personalisation.
- 7) Except for Al Jazeera, each had large amounts of advertising or promoted content.

The decision to use a single webpage from the 9 individual news websites was primarily based on the need to reduce experiment complexity and the number of participants needed and associated costs to obtain a sufficiently high experiment power. A 9x9 within subject incomplete counterbalanced measures design and ANOVA analysis to satisfy four hypothesis is already considered complex.

4.7.1.2 News Articles

As each participant was going to experience nine distorted webpage/article combinations, different news articles had to be used or the experiment would not reflect realistic browsing behaviour. Several steps were taken to reduce the likelihood that the news articles would influence the experiment by acting as a confounding variable. First, emotive topics such as racial issues, religion, immigration, or abortion were all avoided to reduce the likelihood of participant's personal bias influencing the experiment. Second, four news articles were selected from each of the nine news websites being included in the experiment. As the news websites being used in the experiment all belong to serious news agencies or to the quality press, their articles are typically seen as being less biased than other online sources. By selecting articles from these publications, the likelihood of encountering biased language or incendiary claims was reduced. Third, all the articles related to news on the weather but were not weather reports. Thus, providing participants with fresh reading material each time. While biased news relating to the weather is uncommon, it is not impossible, e.g. climate change, man-made climate change, industrial smog. It also ensured that in the extremely unlikely event participants held any personal biases relating to the weather that their effect was common across the experiment. Fourth, the 36 news articles were subject to a pre-test with 12 participants

³¹ Quality press is a term that usually denotes serious print news publications in the UK. It replaced the term broadsheet after many newspapers switched to alternative smaller print formats. Serious news agency describes news organisations which claim to adhere to high standards of journalism, commission original pieces, are often investigative, and claim a duty to inform their readers or viewers of the facts as they are.

using a two-round binary selection process to identify the least biased articles. These were then selected for inclusion in the final experiment.

As this experiment also aimed to investigate the impact of news website category, it was necessary to use news articles with the least amount of bias possible so it would not act as a covariate. Experiments two and three do not investigate news website category and therefore use news articles with a range of bias.

4.7.1.3 Pairing of News Webpages and News Articles

To increase validity, the nine news articles were randomly paired with one of the nine webpages. In some instances, the article and webpage were originally from the same website. These webpage/article combinations were constant throughout the experiment. A copy of each news webpage, including the HTML, CSS and all images, was downloaded. The news article was then inserted into the HTML of each webpage with the associated headings and paragraph tags. Thus, the CSS rendered the webpage/article combination to look exactly like an original from the same website.

Once each distorted webpage/article combination was created, a full-length screenshot of each was taken to be used in the experiment. This removed any possible issues with participant's chosen browsers unequally or differently rendering each distorted webpage/article combination. It also meant that participants could not easily navigate away from the distorted webpage/article combination being tested. It must be noted that a static image of a webpage will never fully replicate the user experience of a live webpage. Differences include loading time of content, animation including advertising, interactive elements such as cursor changes over links, and the ability of user to be able to interact with the webpage and its content such as highlighting text as they are reading it. However, our experiment was focused on the impact of the design, and not the user experience, though they are closely related. It was thus felt that the advantages of a static image over increased realism were worth it.

The decision to pair unbiased news content with branded news webpages was to create as realistic a scenario as possible. If the experiment was set up with no news content or placeholder content such as the Lorem Ipsum³², the participants experience would be unrealistic.

4.7.2 Distortions

To ascertain if any technical features of a news website may affect the perception of bias, each webpage/article combination was subject to eight separate distortions. Each distortion, D0 - D8, involved the removal of the features listed below:

³² <https://www.lipsum.com/>

- D0 - No distortion, the control. An undistorted version of each webpage/article combination to establish a ground truth
- D1 - Removal of branded or explicit advertising, e.g. banner or skyscraper advertising.
- D2 - Removal of reputation lending advertising, e.g. Guardian Dating, Telegraph Wine Club and other products
- D3 - Removal of self-promotional subscription services, e.g. print newspaper or magazine subscription advertisements, premium content advertisements, daily email subscriptions, or campaigns to donate
- D4 - Promoted external content and services, e.g. external clickbait articles from companies like Outbrain and Taboola³³
- D5 - Promoted internal content and services, usually advertisements to lighter style or clickbait type news articles. These are also often provided by the third-party services such as Outbrain and Taboola
- D6 - User interaction facilities, such as website comment facilities or social media comment facilities
- D7 - Article interaction information and buttons: share, send, vote, print, RSS etc.
- D8 - Article embellishments, e.g. author profile info including profile photos, timestamps, location, corrections or clarifications etc.

When a feature was removed from the webpage/article combination, the elements below it was moved up, as if the feature was removed from the underlying code, see Figure 4-2. The features were chosen based on three main factors. First, a feature categorisation review of all nine webpage/article combinations was undertaken to identify common features. Feature categorisation was achieved by applying a common transparent colour overlay to the most obvious common features such as explicit advertising. These included banner ads, search, login, contact facilities, footers etc. Using an iterative process this colour coding was successful in categorising most of the individual elements on each page. Second, their size and prominence were considered, thus advertising features heavily in the list. Third, existing aligned research in the domain which has investigated elements such as the effects of advertising and top level domains on website credibility (Walther et al., 2004), social feedback (Giudice, 2010), seals of approval (Burkell, 2004), and inline or embedded hyperlinks (Borah, 2014).

Webpage footers were also considered for distortion but were ruled out as it was felt that because participants would likely have to scroll to see if they were included or excluded in the webpage design, that it might distort the results of the experiment at it was set up. It should be noted that there is some evidence

³³ www.outbrain.com - www.taboola.com Their software is used on many of the most visited news websites including several tested in this study. It provides implicit news recommendation based on cross website tracking, user models, and multimodal content analysis (Di Massa et al., 2010).

in the literature that users analyse webpage footers for the bona fides of the website sponsor when assessing credibility online (Greer, 2003). Future work may decide to investigate what impact they may have on the perception of bias.

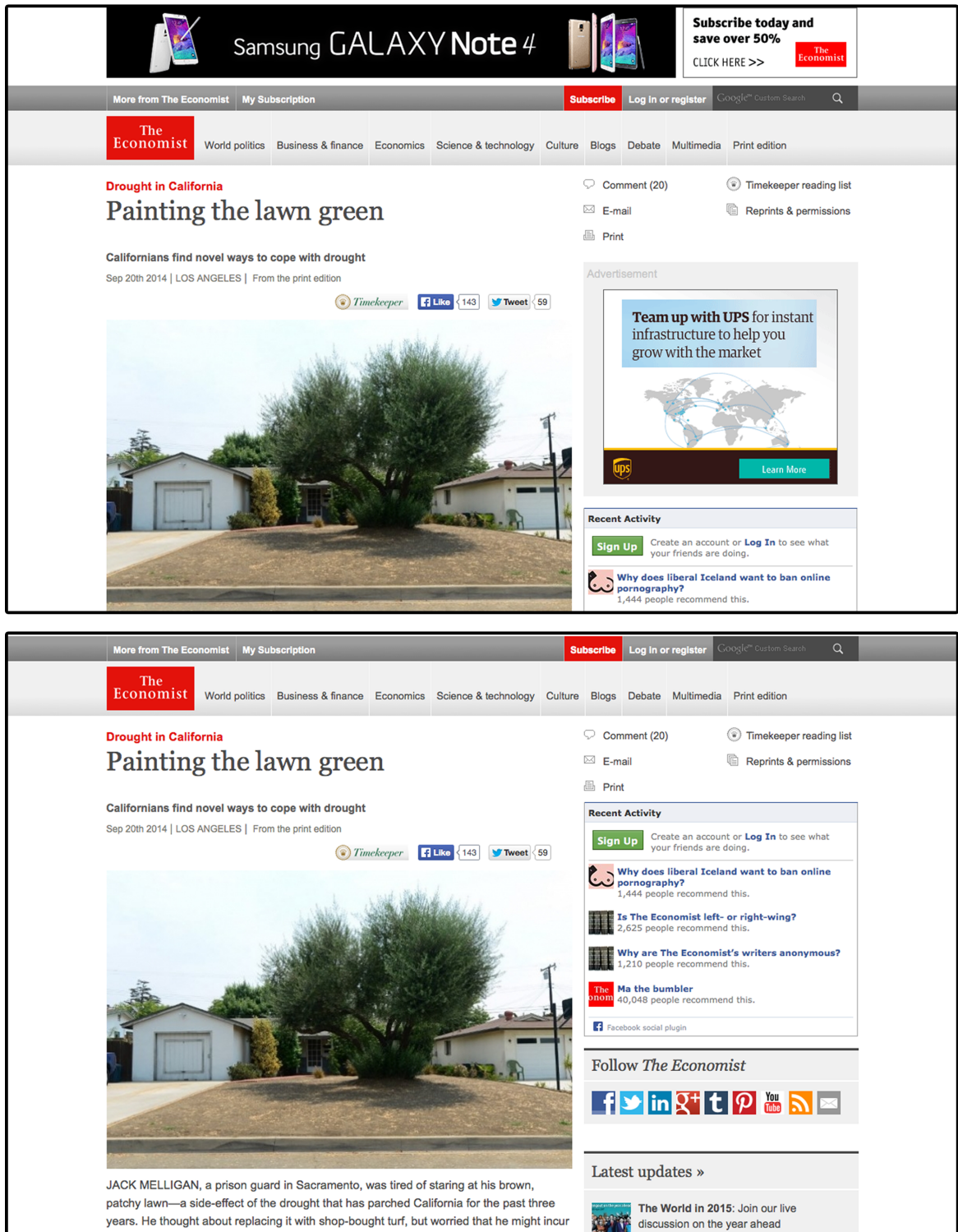


Figure 4-2 Partial screen captures from the Economist website. The top is the D0 (the control), the bottom is the D1 distortion in which all explicit advertising has been removed.

The majority of the ten MTFS detailed in section 2.2.8, and highlighted in Table 2-1, maintain that individual technical or structural features of a webpage allow users to make judgements of credibility, of which bias is a core dimension and measure. These may occur via the peripheral route, heuristic strategy, or the automatic human information process, in the ELM, HSM, and AHIP respectively. In his MAIN model, Sundar staunchly maintains that the underlying technologies behind such features act as a repository of cues, by which the user may judge their affordances and arrive at a credibility judgement (Sundar, 2008). It is the contention of this research that bias in news websites, which has previously, and again subsequently, been shown to be a core dimension of credibility when judging news websites, may also be similarly affected (Fogg, 2003; Spillane et al., 2018).

The decision to investigate eight features (nine webpage/article distortions with D0 control) was primary based on creating a balanced experiment design which would suit ANOVA analysis. The decision to investigate these eight features was primarily based on similar investigation in credibility related research and from identifying common prominent technical and structural features in each of the nine websites.

4.7.2.1 Website Logos and Branding

It was decided not to remove the branding from the websites. First, as it was decided to use real webpages from real news websites, to make the experiment more realistic and any findings more relevant to the domain. Second, it was felt that it would be virtually impossible to hide the identity of the news organisations behind them. Each news agency behind the websites in this experiment spends millions on marketing and their brand to ensure that they are well known or recognisable. Even when removing a name or logo, it was felt that a large proportion of the participants would still likely recognise the websites of e.g. the BBC, the Guardian, or the Economist, even if they did not regularly visit them. Third, as one of the focuses of the experiment was to ascertain whether the category of news website had any effect, it was felt that the logo and branding may contribute to this participant's judgement of such.

4.7.3 The Experiment Platform

The experiment was deployed on an ADAPT Centre server and made public through a reverse proxy to which the crowdsourcing advertisement linked. Crowdsourcing is covered in more detail in section 3.6.2. The experiment website, including the experiment proper was built using the Django Framework deployed on a LAMP stack made public using WSGI. This was the first iteration of the experiment platform that has since been made open-source as part of the CAFE framework. The platform and its underlying technologies are described in more detail in section 3.5 of this thesis. The process of setting up and deploying an experiment measuring credibility or any of its dimensions and measures such as bias using this open-source platform has also been fully documented, which has been included in the GIT repository.

4.7.4 Definition of Bias and Participant Instructions

As has been previously noted, Lowry's analysis of the problem of defining and measuring bias identifies the difficulties, constraints, and the solution: *"The first condition is that one must be willing to accept a relative definition of bias and must recognize that it is impossible to measure bias in human communication in any 'absolute' sense. Therefore, bias in news reports and any other kind of human communication must always be subjectively defined. However, once it is so defined, it may be objectively measured within the context of the definition"* (D. T. Lowry, 1973). Section 2.3.3 of this thesis details the largest categorisation of definitions known in the body of literature. Yet none of these were suitable for this experiment. Firstly, because many such as McQuail's oft cited general definition are meant describe the construct, and are not suited for use in empirical research (McQuail, 1992). Secondly, because many of the definitions are domain specific and focus on media, political, or news bias. Thirdly, they focus on a specific type of bias, such as agenda setting, framing, or coverage bias.

In this experiment, the definition of bias provided to the participants was *"Bias is defined as Positive or negative bias in the treatment of the subject of the news webpage."* During the experiment, participants were provided with instructions and conceptualisations of bias which were designed to make them aware of the role of the visual presentation of the subject material.

- The advertisement placed on the crowdsourcing platform Prolific Academic³⁴ stated that participants were required to rate how positively or negatively biased a series of news webpages were: *"In this study, you will be asked to rate, from -100 to +100, how positively or negatively biased different online news webpages are. -100 being extreme negative bias and +100 being extreme positive bias. Bias is defined as Positive or negative bias in the treatment of the subject of the news webpage."*
- When participants visited the experiment website the Participant Instructions webpage also had similar instructions and definition.
- Participants were also required to complete two instruction tasks to make sure they were aware of how the experiment interface worked. During each of these instruction tasks, a popup window appeared with instructions for what they were to do and the above definition.
- During the actual experiment bias rating task, participants were presented with the following instruction and definition: *"Please rate from -100 to +100 the amount and direction of Bias you perceive in this Design"*, see Figure 4-3 for an example. This was designed make them consider the overall role of the visual presentation of the news article.

³⁴ <http://www.prolific.ac/>

4.7.5 *The Five Stages of the Experiment*

There were five main stages to the experiment; setup (terms and conditions and participant instructions), profile questions, the bias rating tasks, reflective questions, and submit and debrief.

1. Participants were directed to the setup, including the terms and conditions and the instructions stage from Prolific Academic, a crowdsourcing marketplace. Upon agreeing to terms and conditions and viewing participant instructions, they were provided with the definition of bias. They were also provided with two clearly labelled instruction tasks to ensure they understood how the experiment interface worked. Each instruction task had a popup message directing the participants to use the scale beneath each distorted webpage/article combination to rate the amount of positive or negative bias they perceived in each.
2. The second stage focused on eliciting information about the sample of users with questions on physical, political, socioeconomic, and news access habits. This would be used to establish if certain user characteristics impact the perception of bias.
3. The third stage was the bias rating tasks where users were asked to rate the positive or negative bias in nine distorted webpage/article combinations.
4. The fourth stage posed reflective questions on the task they had just completed and on online news and bias in general. Many of the questions were open ended and participants were encouraged to provide further information where possible. These questions and the results can be seen in section 9.20 of the appendix.
5. The fifth and last stage of the experiment provided participants with a final opportunity to agree that their submission be included. They were then finally presented with an experiment debriefing page before they were returned to the crowdsourcing marketplace.

The bias rating experiment (third stage above) was set up as a 9x9 within subject incomplete counterbalanced measures design³⁵. Incomplete counterbalancing was achieved by arranging the nine webpage/article combinations and nine distortions, D0 - D8, in a reduced form Latin square. Participants were randomly assigned to one of 9 diagonal paths through the Latin square that intersected with nine distorted webpage/article combinations, thus ensuring that each participant experienced each webpage/article and each distortion once. To avoid carryover effects or the encroaching effects of task fatigue, once assigned to a path the order of the distorted webpage/article combinations was randomised. The use of Latin squares is further detailed in section 3.6.5 of this thesis.

After rating the perceived bias in three such randomly chosen distorted webpage/article combinations from their assigned path, participants were given the opportunity to compare and re-evaluate their bias ratings for each. To this end, the three distorted webpage/article combinations they had just rated were displayed

³⁵ A detailed description of incomplete counterbalanced measures design and how the experiment framework and platform help researchers to achieve them using Latin Squares and Cubes is provided in section 3.6.5 and in footnote 27

side by side with the bias rating the user had just given them. Participants could then change or fine tune their perceived bias rating. The supposition behind this step was that additional influences on the content or message in an article such as fairness, accuracy or bias may only become apparent when the participant can do a side by side comparison with other comparable news resources. It is these re-evaluated ratings which are used to derive the final results and experiment findings. It would also be possible to see if there are any patterns in the re-evaluation. Examples include the distorted webpage/article combination that was most likely to be re-evaluated, or the most common website to be re-evaluated, and by how much. Two attention questions, to determine continued diligence to the task were also added to each participant's path. Further information on the use of attention questions and other means to increase experiment validity are detailed in section 4.7.7 and 3.6.5 of the thesis.

As a result of this experiment design, when investigating the effects of each distortion on each webpage/article combination the data can be analysed as a series of single variant A/B tests where A is D0 and B is D1 - D8.

4.7.6 Crowdsourcing

Participants for this research were crowdsourced using the Prolific Academic marketplace and paid a fee of GBP£1.25. The experiment was run over two days in late November 2015. Participation was limited to native English speakers based in the U.S, who were over 18 years of age. This was a deliberate tactic. First, to increase the homogeneity of the sample population. Second, because all the news websites and articles were UK based and English language, UK and/or Irish citizens were not recruited as it is more likely they might have pre-conceived opinions or bias against one or more of the publications.

4.7.7 Steps to Ensure Validity of Results

Six steps were taken to ensure the validity of the results. Many of these steps were subsequently incorporated into the CAFE framework and are further detailed in section 3.6.5 of this thesis. They include:

1. Upon beginning the experiment participants undertook two instruction tasks to ensure they fully understood how to use the interface and what was required of them.
2. Two attention questions were added to the start and middle of each participant's path. These were designed to look like the real bias rating tasks so as not to alert participants who were simply clicking through the experiment as quickly as possible to earn their fee. The article for the first attention question detailed the fears of a small community due to toxic bacteria in stagnant flood water full of raw sewage. The article for the second attention question related to sunny weather for the upcoming weekend. Both articles, contained in their webpages, were taken from websites not included in the experiment. The interface for both tasks, including the bias rating scale, was exactly the same as that used in experiment. However, the instruction above each task required the participants to rate whether each article was a positive or a negative news story. Participants were

expected to answer ‘negative’ and ‘positive’ respectively. Contributions from participants who failed one or more of these questions were not assessed.

3. Participants were provided with the opportunity to re-evaluate their bias ratings.
4. Participant’s completion time on Prolific Academic was monitored. Extremely fast (<60 seconds) or quite fast (<10 minutes) completion times would be indicative of either automated bots or participants only interested in completing the experiment as quickly as possible to receive payment.
5. No emotive topics were used.
6. Participants were recruited from a different country to where the websites were based, hopefully reducing the likelihood of any preconceived bias.

A core tenant of the Dual-Process Models of Persuasion is that readers consume information via one of two routes. When the task is new, important, or when the user is engaged, they rely on the central, systematic, or controlled, information processing routes. However, when the user is unengaged, bored, or cognitively lazy, they rely on the peripheral, heuristic, or automatic, information processing routes. Steps 1 - 4 above helped to ensure that participants were engaged in the experiment task, but are not a guarantee that they were engaged in the information processing task, despite the fact that it was part of an experiment and they were getting paid. It is likely that some or even many of the participants took the less cognitively challenging peripheral route. This was expected and suited the experiment aim, as it is more representative of reality.

4.7.8 Statistical Power and Effect Size

When this research was undertaken, there was no guide or comparable experiments to determine appropriate Statistical Power or Effect Size for experiments on perceived bias. There was also very little in the overarching domain of credibility. As is highlighted in sections 2.2.12 and 3.6.1 of this thesis, many studies do not report either of these, despite it being a recommendation of the APA since 1994 (L. Wilkinson, 1999).

As also highlighted in section 3.6.1 of this thesis, it is desirable to have sufficient Statistical Power to detect smaller effect sizes. However, as the complexity of the experiment increases the number of participants required to achieve high Statistical Power also increases. Based on an analysis of related domains it was decided to aim for a Statistical Power to have a 95% chance of detecting an Effect Size of 0.4 or less. An a priori sample size test was computed for ANOVA: Fixed Effects, Special, Main Effects And Interactions using G*Power (Faul et al., 2007, 2009). With the proposed experiment design, at least 151 participants would be required to obtain a 0.4 Statistical Power. Based on Cohen’s determination, an effect of 0.2 is a small, not visible to the naked eye or obvious effect, while an effect of 0.5 is a medium, which is large

enough that the difference is visible or obvious. In total 171 completed submissions were recorded. 9 were rejected due to failing one or both attention questions.

One issue with the design of the experiment was the failure to recognise that a Two-Way Repeated Measures ANOVA requires balanced data to undertake a group comparison. Therefore, only the first 15 submissions to each of the 9 paths (15 is the total number of submissions in the path with the fewest submissions), were included. Thus, 27 submissions were not included in the statistical analysis. Thus, 135 submissions were included in the final dataset from which the findings were derived. G*power was used to determine that this results in a 95% Statistical Power of detecting an Effect Size of 0.42 with an error probability of 0.05%.

Two alternative approaches could have been adopted but were ruled out. First, T-Tests could have been substituted instead of ANOVA, however these do not detect main interactions and do not have the safety of a correction for multiple comparisons such as the Bonferroni or Šidák³⁶ (Abdi, 2007). Second, additional data could have been collected but it would have necessitated running the experiment again several weeks later when the issue was discovered, thus running the risk of an increasingly heterogeneous sample population while delaying the publication of results.

4.7.9 Measuring Perceived Bias with Visual Analogue Scales (VAS)

There is no standard method of measuring bias. Previous attempts to measure perceived bias as one metric of credibility, have used four, five, seven, and ten point Likert scales, (Flanagin & Metzger, 2007; Fogg et al., 2003; Johnson & Kaye, 2002; Kioussis, 2001a; Sundar, 1999). Semantic differential scales and bipolar adjective rating scales have also been used. In some instances, the scales are mono-directional, anchored with derivatives of the absence of bias to extremely biased, while in others they are bi-directional, anchored with terms such as ‘strongly disagree’ / ‘strongly agree’ or ‘describes very well’ / ‘describes very poorly’ (Sundar, 2008). As a result of this lack of cohesion and noting the limitations of Likert scales, such as the lack of suitable and commonly understood terms and the requirement to use non-parametric statistical methods on the resulting ordinal type data, or converting it so that it approaches interval type data, it was decided to adopt a new approach, VAS. The benefits of these were extensively detailed in section 3.6.4 of this thesis. A bi-directional VAS ranging from -100 to +100, anchored with “Negative Bias” and “Positive Bias”, and with large plus and minus symbols, was placed immediately below each distorted webpage/article combination. See Figure 4-3 for an example.

³⁶ Performing a large number of individual tests increases the likelihood of a Type I error, i.e. rejecting the null hypothesis when it is true. Essentially, the larger the number of tests, the easier it is to find a rare event. This is known as inflation of the alpha level. A Bonferroni or Šidák correction for multiple comparisons is designed to reduce the alpha level, thus reducing the likelihood of a Type I error, but also possibly making it harder to detect real effects. The Bonferroni is the more conservative of the two corrections.

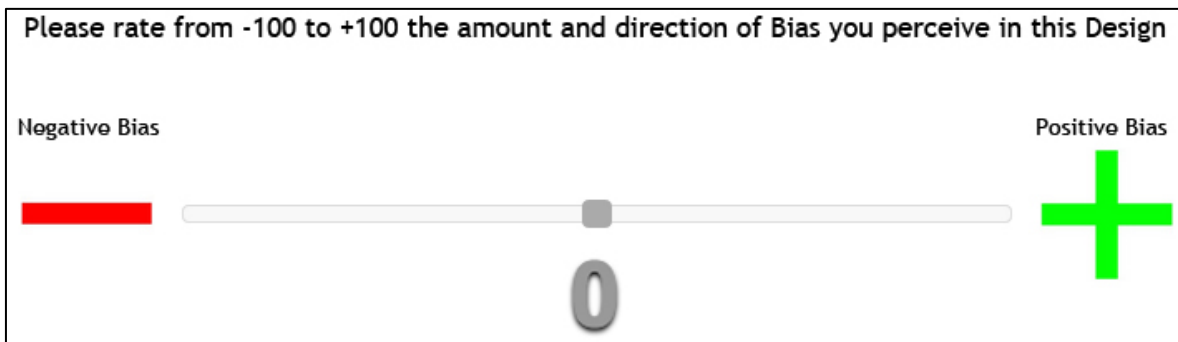


Figure 4-3 Bi-directional Visual Analogue Scale (VAS) from -100 to +100

4.8 Participant Sample Profile Overview

In total 135 submissions were evaluated, 56.3% of which were male, 43.7% female. The average age was 30.1, with 44.4% of respondents between 18 and 25. Participation was limited to the U.S. with 36 states represented. Occupation results point towards a pattern of young students and early career professionals from upper middle-class backgrounds. Only 9.7% of responders reported undertaking skilled, semi-skilled, or unskilled manual labour roles. Professional, semi-professional, and skilled professional accounted for 54.1% of responders. A high rate, 19.3%, of unemployed individuals took part, possibly indicative of a sample with a high proportion of students. Living arrangements also point to a student or young professional profile. 20.7% rent in multiple occupancy, 5.2% in a dormitory and 26.7% live with their family. 14.1% and 12.6% reported owning a home with and without mortgages respectively. Although the average user profile is young, it is representative of the core audience of many news websites. This is backed up by the fact that, in the subsequent reflective questions stages of this experiment, 99.3% of participants reported that they predominantly access news online, see section 9.19.1.2.

4.9 Statistical Analysis and Results - Initial Bias Rating

To increase validity, participants first rated each randomly selected distorted webpage/article combination individually. After three such distorted webpage/article combinations were rated, they were displayed side by side and the participant given the chance to re-evaluate their ratings. This two-stage process was designed to increase the validity of the results and is explained in sections 4.7.5 of this thesis.

The statistical analyses and the results based on the participant's initial bias ratings are presented in section 9.14 of the appendix. The statistical analysis and results based on the participant's re-evaluated bias ratings are presented below in section 4.10. All experiment findings are based on the re-evaluated results shown below. Both sets of data were subject to the same statistical analysis tests and the results are roughly consistent. Where necessary, the below re-evaluated results reference the initial results in the appendix to demonstrate the consistency. The statistical analysis and results based on the initial bias ratings in section

9.14 of the appendix is included for completeness. The frequency of re-evaluation is discussed in section 9.15 of the appendix.

In summary, the results of the Two-Way Repeated Measures ANOVA performed on the initial bias ratings, shown in section 9.14 of the appendix, revealed a statistically significant two-way interaction between the webpage/article combinations and the distortions, $F(40.14, 562.04) = 1.430$, HF $p = .045$. Thus, Simple Main Effects for Webpage/Article Combinations and Simple Main Effects for Distortions was investigated.

Simple Main Effects for Distortions were also performed on the initial bias ratings and the results are shown in Table 9-9 in section 9.14.3 of the appendix. The results show three statistically significant results, D5 for the Spectator, and D7 for the Economist and Reuters. The analysis of the re-evaluated bias ratings, shown below in Table 4-1 in the following section, were largely consistent with the initial bias rating results. It showed the same three results to be statistically significant, along with two additional statistically significant results. These two additional statistically significant results were close to significance in the initial bias ratings.

For completeness, Simple Main Effects for Webpage/Article Combinations were performed on the initial bias ratings and the results are shown in section 9.14.2 of the appendix. However, comparisons of distorted webpage/article combinations between the nine websites do not contribute to the experiment findings or hypothesis. Thus, the results are provided for completeness, but no findings are derived from them.

4.10 Statistical Analysis and Results – Comparative Re-Evaluation Bias Ratings

The following results are derived from the re-evaluated bias ratings, and from which the final experiment findings are also derived.

4.10.1 Main Effects – Comparative Re-Evaluation Bias Ratings

To determine if the re-evaluated results also showed a statistically significant two-way interaction, a Two-Way Repeated Measures ANOVA was carried out on the re-evaluated bias ratings. To conduct the ANOVA non-applicable observations were replaced by their respective webpage/article combination control observations to balance the test.

To undertake a Two-Way Repeated Measures ANOVA, the data and experiment was checked against five assumptions.

- **Assumption One:** The data is a continuous dependent variable.
The data collected from the VAS in this experiment was a continuous dependent variable.

- **Assumption Two:** The experiment has two within-subjects factors, where each within-subjects factor consists of two or more categorical levels.
This has two within-subjects factors, webpage/articles and distortions, and it is a 9x9 design.
- **Assumption Three:** There should be no significant outliers in any cell in the experiment design. An examination of studentized residuals for observations ± 3 standard deviations revealed nine outlier data points. After confirming that these were not data entry or measurement errors, it was decided not to remove the outlier data points as this is considered the option of last resort. It was also decided not to transform the outliers as the data was normally or approximately normally distributed (see point 4 below). It was decided to continue the analysis with the outliers, as comparing the results of ANOVAs with and without the outliers showed little material difference.
- **Assumption Four:** The dependent variable should be approximately normally distributed. A Shapiro-Wilks analysis ($p > .05$) of the studentized residuals and Q-Q plot analysis showed that most of the data observations in each of the distorted webpage/article combination samples were normally distributed with the remainder approximately normally distributed.
- **Assumption Five:** The variance of the differences between levels should be equal. Mauchly's Test of Sphericity (MTS) indicated that the assumption of sphericity had been violated for the two-way interaction. This increases the Type 1 error rate (false positives) for ANOVAs, thus either the Huynh-Feldt (HF) correction or the Greenhouse-Geisser (GG), to produce a more valid critical F-value, should be reported instead. As the estimated epsilon (ϵ) was >0.75 , the more liberal HF correction is reported instead of the more conservative GG.

The Two-Way Repeated Measures ANOVA revealed an increased statistically significant two-way interaction between the webpage/article combinations and the distortions, $F(47.10, 659.43) = 1.473$, HF $p < 0.024$ compared to the initial results, see section 9.14 of the appendix. Thus, Simple Main Effects for Webpage/Article Combinations, and Simple Main Effects for Distortions were once again undertaken on the re-evaluated results.

4.10.2 Simple Main Effects for Distortions – Comparative Re-Evaluation Bias Ratings

Simple Main Effects for Distortions in the form of One-Way Repeated Measures ANOVAs were conducted on the re-evaluated bias ratings comparing the respective controls of each webpage/article combination (D0) to its respective distorted webpage/article combinations (D1-8). In Statistical Package for the Social Sciences (SPSS)³⁷, running Simple Main Effects for a Two-Way Repeated Measures ANOVA is the same as running separate One-Way Repeated Measures ANOVA. The intersecting cell of each webpage/article x distortion combination in Table 4-1 shows the Mean bias ratings, Standard Error, F-Statistic and Significance value. The degrees of freedom (DF) for the factor and error is (1, 14) and Confidence Interval

³⁷ <https://www.ibm.com/ie-en/products/spss-statistics>

(CI) is 95%. D0 is the control with no distortion applied. The cells in columns D1 – D8 show the results of the Simple Main Effects comparing each to their respective D0 controls. Scores closest to zero indicate less biased webpage/article combinations. The five highlighted cells indicate the five significant results found in the statistical analysis of the re-evaluated data. To demonstrate the consistency in the findings between the initial and re-evaluated analysis (see Table 9-9 in appendix 9.14), the three light grey cells indicate where significant results were also found in the initial bias ratings. The dark grey cells indicate where significant results were only found in the re-evaluated statistical analysis. N/A indicates the feature was not present.

The results of the One-Way Repeated Measures ANOVAs on the re-evaluated bias ratings for the distorted webpage/article combinations confirmed the results of the initial bias ratings presented in section 9.14 of the appendix, and revealed three additional significant interactions. The re-confirmed initial significant results are highlighted in light grey, while the additional significant results are highlighted in dark grey, see Table 4-1. Two distortions have a statistically significant effect on four of the nine websites involved. D5 on The Spectator and The Telegraph, and D7 on The Economist, Reuters, and The Spectator. There is, however, variation in directionality of the influence.

4.10.2.1 D5

In the initial bias rating results shown in Table 9-9 in the appendix, D5 significantly increased the perception of negative bias on the Spectator D5 webpage/article combination compared to its D0 control. This same effect was also shown in the re-evaluated results shown in Table 4-1. The Spectators D5 re-evaluated bias rating (M -37.667) is significantly more negative than its D0 control (M -7.53). There is also only a slight variation in the results from the initial D5 rating (M-37.73) and the re-evaluated results (M -37.667).

The re-evaluated bias ratings revealed an additional significant finding for the D5 distortion. It was shown to significantly increase the perception of positive bias on the D5 Telegraph webpage/article combination (M 20.20) compared to its D0 (M 0.60), a mean difference of 19.60. It should be noted that this effect was nearly significant in the initial bias rating data with a p value of 0.58. It is possible that the different category of news website was responsible for this inconsistency of effect.

Table 4-1 Results of re-evaluated One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions test using the Bonferroni correction for multiple comparisons.

	D0	D1	D2	D3	D4	D5	D6	D7	D8
Guardian	M -1.20 SE 6.977	M -8.333 SE 6.508 F .671 p .426	M -2.133 SE 7.899 F .006 p .940	M 5.20 SE 7.982 F .353 p .562	M 9.333 SE 5.681 F 1.087 p .315	M 2.00 SE 11.689 F .083 p .778	M 15.133 SE 9.166 F 2.165 p .163	M 11.00 SE 6.590 F 1.955 p .184	M -9.200 SE 6.515 F .920 p .354
Telegraph	M 0.60 SE 5.623	M -5.933 SE 4.173 F .696 p .418	M -2.47 SE 5.073 F .145 p .710	N/A	M -0.53 SE 7.866 F .019 p .891	M 20.20 SE 7.563 F 4.881 p .044	N/A	M -13.00 SE 7.764 F 2.023 p .177	M 2.467 SE 6.803 F .033 p .858
Independent	M 5.133 SE 7.560	M -16.467 SE 9.937 F 2.479 p .139	M -19.133 SE 10.441 F 3.576 p .079	N/A	M 0.667 SE 10.129 F .148 p .707	N/A	M 16.067 SE 10.094 F .852 p .372	M 1.533 SE 7.941 F .100 p .757	M -5.733 SE 7.651 F .828 p .378
Economist	M 12.13 SE 10.154	M 30.667 SE 6.563 F 3.159 p .97	M 20.133 SE 8.065 F .279 p .606	M 20.00 SE 9.623 F .341 p .568	N/A	M 5.867 SE 6.828 F .253 p .623	M 5.80 SE 9.739 F .874 p .366	M 36.13 SE 8.216 F 7.378 p .017	M 11.933 SE 8.867 F .000 p .988
Spectator	M -7.53 SE 9.344	M -29.93 SE 13.123 F 2.393 p .144	M -20.73 SE 8.925 F .925 p .353	M -16.267 SE 7.879 F .583 p .458	N/A	M -37.667 SE 7.678 F 7.154 p .018	M -4.40 SE 11.475 F .035 p .853	M -33.00 SE 9.571 F 5.564 p .033	M -13.87 SE 7.665 F .283 p .603
New Statesman	M 22.27 SE 11.608	M 16.07 SE 9.375 F .161 p .695	N/A	M 21.60 SE 9.219 F .002 p .967	N/A	M 24.60 SE 8.448 F .045 p .836	M 29.60 SE 8.784 F .382 p .546	M 12.867 SE 9.210 F .604 p .450	M 29.80 SE 7.633 F .325 p .578
Al Jazeera	M -0.27 SE 5.483	N/A	N/A	N/A	N/A	M 7.87 SE 7.802 F .877 p .365	M 17.47 SE 8.927 F 2.641 p .126	M 7.80 SE 5.560 F 1.971 p .182	M 2.533 SE 8.555 F .058 p .813
BBC	M 6.00 SE 7.114	M 21.200 SE 10.105 F 1.518 p .238	N/A	N/A	M 14.667 SE 8.617 F .490 p .496	M 21.000 SE 8.617 F .490 p .496	N/A	M 15.40 SE 7.792 F .951 p .346	M 10.33 SE 4.383 F .361 p .558
Reuters	M -7.733 SE 4.318	M -21.133 SE 7.204 F 2.498 p .136	N/A	N/A	M 10.667 SE 9.782 F 2.663 p .127	N/A	N/A	M 12.40 SE 6.129 F 7.158 p .018	M 7.667 SE 4.318 F 1594 p .227

4.10.2.2 D7

In the initial bias rating results shown in Table 9-9, D7 significantly increased the perception of positive bias for two webpage/article combinations, the Economist and Reuters. Analysis of the re-evaluated bias ratings shown in Table 4-1 confirmed these initial findings and showed only marginal variation in the perceived bias in the Reuters D7 (M 12.60 to M 12.40) while the Economist (M 36.13 to M 36.13) was the same. The re-evaluated results also revealed an additional significant interaction for D7 on the Spectator. The re-evaluated results showed the D7 distortion significantly increased the perception of negative bias (M -7.53 to M -33.00). This difference was close to being significant in the initial results.

4.10.3 Simple Main Effects for Webpage/Article Combinations – Comparative Re-Evaluation Bias Ratings

Simple Main Effects for Webpage/Article Combinations was also conducted on the re-evaluated bias ratings. All Confidence Intervals (CI) are 95%. All tests passed MTS except for D0 and D1, where HF corrections are reported as the estimated epsilon (ϵ) >0.75. Essentially, these are results of comparisons of the distorted webpage/article combinations between websites. E.g. comparing all nine of the D0 webpage/article distortions shown in Table 9-9 to each other. It should be noted that none of the experiment findings or the experiment hypothesis rely on the results of the Simple Main Effects for Webpage/Article Combinations reported in this section. They are reported in section 9.15 of the appendix for completeness.

4.10.4 Most Positively and Negatively Biased

At the end of the experiment, participants were shown large thumbnails in a 3x3 arrangement of all 9 of the distorted webpage/article combinations they had rated. They were then asked to select which webpage/article combinations they perceived to be the most positively biased. This was then repeated to select the most negatively biased. The results are shown in Figure 4-4.

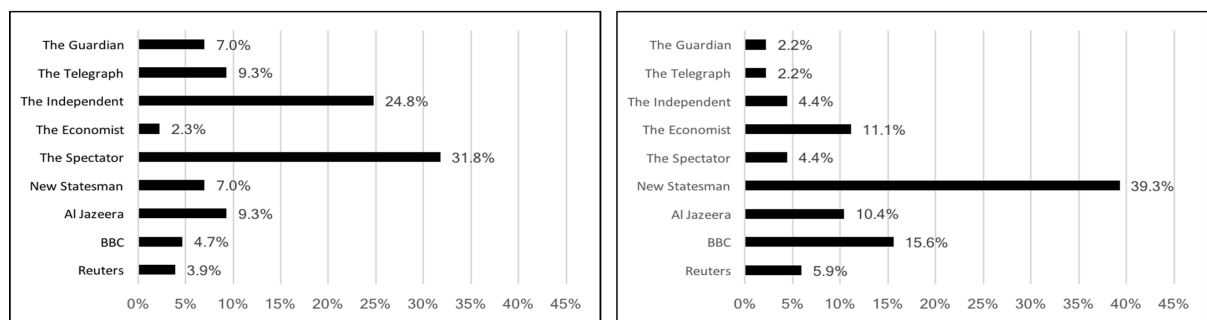


Figure 4-4 The most negative on the left – the most positive on the right.

A very high proportion of participants, 31.8% and 24.8% selected distorted webpage/article combinations from The Spectator and The Independent as the most negatively biased webpages. This correlates strongly with the previous results shown in the initial bias ratings shown in Table 9-9 and the re-evaluated bias ratings shown in Table 4-1. Distorted webpage/article combinations from the same websites were each selected by just 4.4% of participants when asked to select the most positively biased.

The opposite is true of the New Statesman. Distorted webpage/article combinations from this website were selected by 39% of participants as the most positively biased, and by only 7% as the most negatively biased. The results of this also correlate strongly with the results in the initial bias ratings shown in Table 9-9 and the re-evaluated bias ratings shown in Table 4-1. Considering the strength of these results and the fact that they are backed up by the results shown in both tables, the underlying design and/or the overall visual presentation or aesthetic is the most probable cause for such positive and negative sentiment. Just as

previous work by Robins and Holmes has demonstrated a positive correlation between levels of aesthetic treatment and credibility, it appears that the visual presentation of these websites is influencing users' visceral perception of bias (Robins & Holmes, 2008). It should also be noted that previous work in the domain by Lindgaard et al. has demonstrated that users make significant judgements about webpages within just 50ms (Lindgaard et al., 2006).

4.11 The Impact of News Website Category on The Perception of Bias

4.11.1 Traditional Print News Agencies

Participants were also influenced by the category of webpage/article combination they were rating. The webpage/article combinations of the traditional print newspapers were less likely to have a strong effect on the perception of positive or negative bias compared to those of news magazines or international multi-format news agencies, see Table 4-1. This is most noticeable in the results for the Guardian and the Telegraph. It is also noticeable, though more inconsistent in the Independent. With NA values replaced by those either side, the average bias rating for all the traditional print webpage/article combinations was just 0.67.

4.11.2 News Magazines

As can be seen in Table 4-1, the perceived bias rating for the webpage/article combinations from the news magazines, the Economist, The Spectator, and the New Statesman, elicited much more polarising bias ratings than those of traditional print publications. The most obvious manifestation of this polarisation is the overtly positive bias ratings the Economist and the New Statesman received, almost always the most positive, and the extreme negative bias ratings the Spectator received, see Table 4-1. These findings are backed up by the results shown in section 4.10.4 and Figure 4-4, which showed that most users selected the New Statesman as the most positive and the Spectator as the most negative when presented with all the distorted webpage/article combinations they had just rated.

4.11.3 International Multi-Format News Agencies

The perceived bias in the distorted webpage/article combinations belonging to the websites of international multi-format news agencies is much more varied. Reuters webpage/article combinations were seen as both positively and negatively biased, those belonging to the BBC were seen as positively biased, while those belonging to Al Jazeera were among the least evocative. This variation is possibly due to the large disparity between the design and visual presentation of these websites.

4.11.4 The Impact of the Overall Design and Visual Aesthetics of a News Webpage on the Perception of Bias

The findings of the study presented in this chapter demonstrate that the category of news website has an impact on perceived positive and negative bias. Confirming this are the findings presented in section 4.10.4 and in Figure 4-4, where participants were asked at the end of the bias rating stage of the experiment to pick the most positively biased and the most negatively biased distorted webpage/article combinations. Out of the nine, 39.3% of participants selected the New Statesman as the most positive overall, while 31.8% selected The Spectator as the most negative. These results closely correlate to the ANOVA results shown in Table 4-1, detailing the impact of distortions investigating technical features of the webpage. The results show that participants had an overwhelming perception of positive bias in the New Statesman, which remained even as each distortion was applied. The Spectator, on the other hand received negative bias ratings on the control and on each distortion. Just as interestingly, certain webpages did not incite strong feelings of bias. Reuters was selected as the most positively biased and negatively biased by just 5.9% and 3.9% of participants respectively. The Guardian and the Telegraph had similarly low scores.

4.12 The Impact of User Characteristics on the Perception of Bias

The research revealed that participants had an overall tendency to perceive a slight positive bias in the distorted webpage/article combinations tested. $N = 930$, $M = 4.31$, $SD = 34.91$.

4.12.1 Physical User Characteristics

Research on user characteristics such as age, sex, education, Internet usage, involvement etc. have a long involvement in credibility research. A detailed breakdown of such studies can be found in the review by Metzger et al. (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003). To the best of the authors knowledge, this is the first time such research has been done on bias.

4.12.1.1 Gender

The tendency to perceive a slight positive bias in the distorted webpage/article combinations was slightly more noticeable among women $N = 407$, $M = 5.03$, $SD = 34.10$ than men, $N = 523$, $M = 3.74$, $SD = 35.54$.

4.12.1.2 Age

The Mean bias ratings of participant's broken down by age demonstrates that the younger a participant was, the more likely they are to perceive a positive bias. A two-tailed Z-Test, ± 1.960 , conducted on the Mean bias rating for each age bracket revealed a significantly more positive perception of bias for those in the 18 to 20 age bracket $M = 11.16$, $z = 2.02$. While this is a smaller category than the others, a clear trend between youth and perception of positive bias is evident in the results shown in Table 4-2.

Table 4-2 Impact of participant's age on the perception of bias.

Age	% of 930	Mean Bias Rating	SD
18 – 20	6.2	11.16*	33.74
21 – 30	54.3	5.32	37.16
31 – 40	24.0	2.01	27.83
41 – 50	7.8	2.07	42.70
51 – 60	4.6	1.28	30.38
61 - 70	3.0	0.68	30.75

4.12.2 Education, Income, and Socioeconomic

Education, Occupation, and Living arrangements were shown to have no impact on perceived bias. The sample profile breakdown for each and the bias ratings are shown in section 9.17 of the appendix.

4.12.2.1 Income

The income range reflects the age profile of the participants. 29.6% earned less than \$10,000 per annum and 43% earned less than \$20,000. The average wage of the participants in the survey was \$33,622.47. This is in comparison to an average wage in the U.S. of \$46,481 in 2014 (Social Security, 2016). The low average wage is likely due to the high response rate from young participants, especially college age responders. This is also backed up by the reported high rate of participation in education. Most responders also had a low level of disposable income. 11.1% reported not being able to save any money each month and 51.8% report saving ten per cent or less each month.

Overall income levels show some inconsistency as indicators of bias. A two-tailed Z-Test, ± 1.960 , conducted on the Mean bias rating for each income bracket revealed a significantly more negative perception of bias for those in the \$70,000 to \$79,999 pay bracket, $M = -22.75$, $z = -1.99$ and for those in the \$90,000 to \$99,999 pay bracket, $M = -22.50$, $z = -1.97$. However, participants in the \$80,000 to \$89,999 pay bracket show normal levels of perceived bias. Consequently, further testing is required in future to verify this outcome, see Table 4-3.

Table 4-3 Income levels and the perception of bias.

Income	%	Mean Bias Rating	SD
\$10,000 or less	30.0	7.48	34.03
\$10,000 - \$19,999	13.3	4.60	41.30
\$20,000 - \$29,999	10.0	0.52	31.51
\$30,000 - \$39,999	15.6	5.51	33.15
\$40,000 - \$49,999	9.4	2.47	33.84
\$50,000 - \$59,999	7.3	1.85	28.74
\$60,000 - \$69,999	5.3	0.94	34.66
\$70,000 - \$79,999	0.9	-22.75*	47.20
\$80,000 - \$89,999	2.5	1.61	27.21
\$90,000 - \$99,999	0.6	-22.50*	40.98
\$100,000 or more	5.2	4.42	42.32

4.12.3 Political Ideology and Affiliation

Participants were also asked about which elections they vote in, their political leaning, and ideological viewpoints. The population sample was much more liberal/democratic than conservative/republican with a strong belief in voting. For brevity, only the significant results are shown below. The rest of the data is presented in section 9.18 of the appendix.

4.12.3.1 Party Affiliation

Combined, those identifying as Democrat and Republican had a positive Mean bias ratings of N = 45.5%, M = 4.98, SD = 34.08 and N = 16.1%, M = 2.59, SD 33.31. Like the political ideology results those identifying as “Other” and “Don’t Know” had raised positive perceived bias. A two-tailed Z-Test, ±1.960 showed that the above average positive perception of bias by those who answered “Don’t Know” was significant, M = 17.35, z = 2.11, see Table 4-4. The strong positive bias perceived by those who answered “Other” and “Don’t Know” for political ideology and political party affinity is the result of their likely affinity for alternative political movements in the US such as the Green Party. It is possible that the positive bias perceived in the news stories relating to the weather area result of their personal believes about nature, climate change and the importance of the environment.

Table 4-4 Political party affinity and perception of bias.

Political Party Affinity	%	Mean Bias Rating	SD
Strong Democrat	13.0	6.43	36.25
Not so Strong Democrat	12.7	5.84	35.82
Independent Leaning Democrat	19.8	3.47	31.49
Independent	27.4	0.71	32.42
Independent Leaning Republican	4.6	0.33	33.15
Not so Strong Republican	7.1	7.71	33.50
Strong Republican	4.4	-3.27	32.71
Other	4.3	6.40	39.26
Don’t Know	6.7	17.35*	46.99

4.12.4 Internet Usage and News Access Habits

Participants were also asked a range of questions to establish their Internet usage and news access habits to establish if their impacted their perception of bias. No significant results were found. This data is presented in section 9.19 of the appendix.

4.12.5 Summary of User Characteristics Results

Until this research was undertaken, there was no information as to whether user characteristics had any impact on the perception of bias in online news. Consequently, an exploratory approach was adopted, and a wide range of questions were included in the experiment. User characteristics showed mixed results as an indicator of positive or negative bias in the webpage/articles tested.

The most consistent indicator is age. As expected, younger participants were more likely to rate the distorted webpage/article combinations as positively biased, particularly those under the age of eighteen. It should be noted however that this group is a particularly small sub-section of participants, see Table 4-2. Nevertheless, there is an overall perceptible positive to negative slant as participants get older. Education level also shows that those who have achieved higher qualifications, likely older responders given the age profile of the group, were also more likely to perceive a negative bias, see Table 9-12. A more conservative position is off the natural tendency of those who have reached a high socioeconomic status due to a desire to protect their position. The results also show that those who report working in professional jobs are also more likely to perceive positive bias, see Table 9-13. However, income level proved a contradictory, and an inconsistent indicator. The results shown in Table 4-3 show two high income categories have significantly negative perceptions of bias. Usually, but not always, professional jobs indicate a higher income. It should be noted that as the income categories are arbitrary, a slight change might have resulted in more consistent negative perception of bias for those reporting high incomes. It should also be noted that very few participants earned such high wages. Consequently, it is not impossible that the two significant results are the result of an anomaly in the data or an unlikely collection of outliers. Further research is therefore recommended. Political ideology and political party affinity also provided mixed results. However, interestingly those identifying as “Other” and “Don’t Know” perceive a strong positive bias, see Table 9-14 and Table 4-4. It is possible that those who do not affiliate with an established political position are more likely to access their news from alternative sources and have a low opinion of established news media, thus resulting in a higher rate of perceived bias. However, this would require further study.

Overall, user characteristics are inconsistent predictors of perceived bias, though there are very definite trends visible in the data. Thus, further research is required to confirm these findings and to see if any other characteristics may influence how users perceive positive and negative bias in the news. These findings should act as a foundation and a guide to designing more targeted studies. They also serve to provide an in-depth breakdown of the experiment sample.

4.13 Hypothesis Results

4.13.1 Primary Hypothesis Results

H₀ Perceived bias in the news articles will not be affected due to distortions in the news webpage.

H_A Perceived bias in the news articles will be affected due to distortions in the news webpage.

Based on the evidence presented in sections 4.10, and 4.11, it is possible to reject the null hypothesis and accept the alternative hypothesis, perceived bias in the news articles is affected due to distortions in the news webpage.

This demonstrates, that users' perception of bias in a news article is affected by its visual presentation. This effect has been previously observed, where the visual presentation of information online is known to impact perceived credibility, of which bias is a core dimension and measure when judging information online.

4.13.2 Secondary Hypotheses Results

4.13.2.1 Individual Features

H₀ Perceived bias in the news articles will not be affected due to removing individual features of the webpage.

H_A Perceived bias in the news articles will be affected due to removing individual features of the webpage.

The inconsistency of the evidence presented in section 4.10 means that the experiment fails to reject the null hypothesis. While two distortions, D5 and D7, were shown to have a significant impact on the perception of bias in the initial bias ratings and in the re-evaluated bias ratings shown in Table 9-9 and Table 4-1, and even though a number of steps were taken to ensure the validity of the results, the inconsistency of the effect means that it cannot be categorically rejected without risking a Type 1 error.

However, the results from this experiment mean that this warrants further investigation. It is possible that for the effects to be more consistent and/or visible, that multiple distortions should be combined.

4.13.2.2 User Characteristics

H₀ User characteristics will not affect perceived bias in the news articles.

H_A User characteristics will affect perceived bias in the news articles.

While there were some noticeable trends evident, particularly age and education level, and several significant results, their inconsistency means that the experiment fails to reject the null hypothesis and does not accept the alternative hypothesis.

It was also surprising that participants' political leaning and ideology, and their news access habits, failed to reveal any consistent conservative or liberal patterns, yet those who "Other" or who "Do Not Know", revealed consistent patterns in their bias ratings.

4.13.2.3 Category of News Website

H₀ Perceived bias in the news articles will not be affected due to the category of news website.

H_A Perceived bias in the news articles will be affected due to the category of news website.

Due to the inconsistency in the results across the three categories of news websites, the experiment fails to reject the null hypothesis.

However, if the hypothesis was to be judged on the webpage/article combinations of print newspapers only, the null hypothesis would be rejected and the alternative accepted. This is due to the results showing that the websites of traditional news mediums are considered less biased than those of news magazines or international multi-format news agencies. This trend was consistent across the majority of webpage/article combinations tested and was more evident in their averaged results. As the articles were all rated as unbiased in the pre-test, the most likely explanation is the different styles of their design and visual presentation.

One note of caution, and a possible confounding variable, is the possible impact of the brand and name of each news website. As discussed in section 4.7.2.1, it was decided not to remove the branding from the experiment webpages to make the experiment more realistic and because it was felt that simply removing the branding would not hide their identity, and making more substantive changes would impact the realistic nature of the experiment. Consequently, it was decided to address this possible confounding variable in subsequent experiments.

4.14 Reflective Questions

The third stage of the experiment involved the participants answering a series of reflective questions based on the bias rating task they had just completed. It also involved asking them a broad range of exploratory questions on bias itself, how they identify it, how it is communicated to the user, and how they react to it.

Many of the questions encouraged the participants to fill out open comment boxes so that they were free to add additional information. This purely exploratory approach was adopted to better inform the design and development of subsequent experiments. The reflective questions, their format, and the answers are recorded in section 9.20 of the appendix. For brevity, the following sections only provides an overview of the major findings. More detailed information on participant's responses can be found in section 9.20 of the appendix.

4.14.1 Least Biased Medium

When asked to choose the least biased medium, surprisingly online was considered almost as unbiased as traditional print newspapers, with 39.26% compared to 40% of participants selecting it, see Table 9-17 in section 9.20.1 of the appendix. Just 14.07% named radio and 6.66% named television as the least biased mediums. It is possible that due to the huge range of news resources available online that the medium as a whole, is considered nearly as unbiased as print newspapers, which have traditionally been seen as the least biased. This likely contributed the increasing reliance on the medium as a primary source of news among the participants, see Table 9-15.

4.14.2 Most Noticed Forms of Bias

Participants were also asked how often they notice seven common forms of bias in the news. The detailed results of these questions can be found in the tables in section 9.20.3 in the appendix, a summary of the results is presented below in Table 4-5. Three interesting facts are evident from this data. First, compared to the other forms, the sample group did not highlight gender or health care as topics which receive especially biased treatment in the press. Second, 88.89% of participants claim to always or sometimes notice political bias in the news they consume. Third, the dichotomy in the results of political party favouritism, where 26.67% of participants claims to always notice bias in favour of one party or another while over half, 52.59% claim to never notice it. These results are somewhat at odds with the results of how often they notice political bias.

Table 4-5 Regularity of biased reporting in the news on common divisive issues.

	Gender	Political	Racial	Climate Change	Health Care	Religious	Political Party Favouritism
Always	14.81%	43.7%	20.74%	25.19%	13.33%	31.11%	26.67%
Sometimes	38.52%	45.19%	48.15%	48.89%	42.96%	42.96%	12.59%
Rarely	37.77%	9.63%	22.96%	20.74%	33.33%	18.52%	8.15%
Never	8.88%	1.48%	8.15%	5.19%	10.37%	7.41%	52.59%

Participants were also asked via multiple choice, which political topics regularly received biased coverage, see Table 9-25. Topics of national importance and in particular elections, are much more likely to receive biased coverage than state or local issues. In the open comments section, participants also listed additional topics such as abortion, birth control, gun control, science, and global or geopolitical issues.

4.14.3 Common Features and Characteristics of a News Webpage, Including the News Article, which can Impact or Convey Bias

In a series of questions, participants were asked which common features, or characteristics of a news website's design, reflected in a news webpage, including the news article contained therein, convey bias to the reader. The results of these questions are contained in the tables in section 9.20.4 in the appendix, while the results are summarized below.

4.14.3.1 News Article Features Impact on Perceived Bias

Table 9-26 highlights which features within a news article, participants perceived as commonly used to bias or convey bias. Three features unsurprisingly stood out from the rest. They include: News Article (Content or language), News Article Headlines (Content or language), and News Article Photographs.

4.14.3.2 News Website Homepage Features Impact on Perceived Bias

Table 9-27 highlights which features of a news website's homepage that participants believe commonly convey or add bias. As expected, 93.33% and 76.30% selected two options relating to Homepage

Headlines, Leads and Links; Word choice, language, word difficulty, grammar, and Space, coverage, size, position or prominence. However, large proportions, 85.93% and 82.22% also selected two options relating to the overall combination, spread, or grouping of news stories, headlines and images.

Somewhat unsurprisingly, the features of a news website's homepage were considered much more prone to bias than many of the same or similar features of an individual webpage. This is likely due to the fact that they are designed to attract attention as readers are essentially: "*shoppers for headlines*" (Tannenbaum, 1953). While news headlines have been the focus of a range of studies into news bias, they have predominantly focused on textual content and to a lesser degree their size and prominence. Several examples of studies can be seen in Table 9-7 and in Table 2-3. Section 2.3.5.1 of the State of the Art chapter of this thesis also contains an in-depth analysis several of these studies. However, to the best of the author's knowledge, there has been no work undertaken on the combination of text, images, link, and article lead or teaser content, which combined, form the headlines on news websites homepages. Participants in this study have shown that in their opinion they are among the worst offenders for conveying bias. It is possible that there is some acceptance within journalism that news headlines are somewhat exempt from normal rules of quality or serious journalism. This was highlighted as early as 1946 by Kriesberg who when writing about unwarranted headlines states: "*there is no attempt to judge headlines by the criterion of 'true' news.*" (Kriesberg, 1946). Based on participants' responses, bias in the headlines on new websites homepages warrants further investigation.

4.14.3.3 Characteristics of a News Website's Design and Visual Presentation Impact on Perceived Bias

Participants were also asked whether certain characteristics of a news website's design might reduce, increase, or have no effect on their perception of bias. As defined in section 2.2.11, characteristics are the overarching themes behind a website's design which the designer is trying to impress on the user as part of the overall brand of the organisation behind the website. Table 9-28 in the appendix contains detailed results.

Participants were presented with a thirty-eight-item list which included common characteristics which news agencies, online editors, and news website designers want to convey to consumers, such as Modern, Dynamic, and Trustworthiness. The list also contained items which designers typically do not want to convey to users such as Cheapness, Clashing, or Bad Layout.

These items were selected from the literature investigating characteristics of successful websites (Tarafdar & Zhang, 2005; Zhang et al., 2000), including those specific to news websites (Aranyi & Van Schaik, 2015; Hope & Li, 2003), the characteristics of online trust, an aligned and important domain (Seckler et al., 2015), and website satisfiers and dissatisfiers (Zhang & von Dran, 2000). Other influences include the scales and variables used by Kenix on her work on commercialism and the visual identity of alternative

and mainstream news websites (Kenix, 2013). Other items were selected from the literature measuring website quality and credibility (Barnes & Vidgen, 2000; Fogg et al., 2003).

From the user submissions, four characteristics stood out from the rest which they believe reduce perceived bias; Seriousness, Professional, High Quality Aesthetics, and Good Design. Two characteristics; Tabloid and Gaudiness were unsurprisingly found to increase perceived bias. Tabloid was included as it is a recognised news website visual style.

The impact of design and visual presentation on perceived bias in news websites is a multifaceted problem. As the State of the Art demonstrated, there is a long history of research on the impact of the visual presentation of news through its dissemination in traditional print and television mediums, yet there has been very little work on the impact of the visual presentation during the dissemination of news online. The design and visual aesthetics of a news website are a complex combination of underlying software supporting individual technical features and affordances, overlaid by a common design, and constantly churning content. Consequently, it was necessary to break the problem down to tackle individual aspects of it separately. While the main part of the first experiment was designed to investigate the impact of individual features of a news website's design, which are reflected in a news webpage's aesthetic, on the perception of bias, the reflective questions were designed to inform the design and development of the second and third experiment. It was clear from the users' answers that the characteristics of a news website's design, which are also reflected in a news webpage's aesthetic, have a large impact on their perception of bias in the news articles they contain.

A second and unintended benefit of the results of this question is its usefulness for designing a future experiment to encapsulate and measure the concept of bias. Currently no research exists, such as that done by Gaziano and McGrath, to determine the individual dimensions of the concept of bias in order that it may be measured more accurately in future (Cecilie Gaziano, 1988).

4.14.3.4 The Impact of a News Website's Content Focus on Perceived Bias

Participants were also asked whether the focus or the type of content a news website typically displays impacts their perception of bias in the news articles it contains. The results which are in Table 9-29 in the appendix show that participants found that the salaciousness and sordidness of content, and the invasiveness of pictures were all likely to increase their perception of bias. These factors, and others including a focus on celebrity news and pictures with a high proportion of bare skin or which could be considered sexual, are all predominantly found in tabloid style news websites. These findings back up the results of the previous reflective question which found that if a news website had the characteristics of a tabloid that it would be considered more biased. Surprisingly, focus on sex was considered by roughly equal proportions of the sample group to reduce, increase, or have no effect on the perception of bias.

This question was asked as there is currently no research known to the author on whether the surrounding content or the visually obvious focus of a news website might impact the perception of bias within a news article. The supposition behind this is that as a user moves through a news website, they may form opinions as to how a news agency might report a news story based on a passing visual analysis of the type and focus of their news articles. This is similar to the effect of “*information scent*”, where the surrounding content and technical features of a webpage, including supporting content, personalised recommendations, and related or trending content, emit information scent. Users use these visual cues to make heuristic judgements about the acceptability and credibility of the information before them. In the words of Sundar, they: “*emit ‘information scent’ helpful in making quick decisions about the quality of the information available for consumption.*” (Sundar, 2008; Sundar et al., 2007).

While these results proved very interesting, it was decided that it would not be the focus of the subsequent experiments as it would move the research closer to analysis of the textual content rather than its primary aim of the impact of the visual presentation. However, this will likely form an avenue for future work.

4.14.3.4.1 The Impact of Website and Third-Party Functionality on Perceived Bias

Participants were also asked a series of questions on the impact of website functionality, third party functionality, personalisation, and advertising on the perception of bias. The results of which are contained in the tables in section 9.20.6 in the appendix.

A large proportion of participants claim to have detected bias in trending content, comment facilities, and ‘also read’ recommended content facilities, see Table 9-30 Chakaborty et al., has previously demonstrated that such facilities may contribute to coverage bias by limiting the news articles a user is exposed to in favour of recent or trending content (Chakraborty et al., 2015). Participants were asked whether they had ever detected bias in third party social media facilities or in personalisation facilities, see Table 9-31 and Table 9-32. However, the results were for the most part inconclusive. Lastly, participants were also asked whether they have detected bias in a range of advertising, see Table 9-33. The results of this indicate that a large proportion have detected bias in a broad range of common website advertising formats and this likely warrants further investigation in future.

These questions were asked as there is little or no information currently as to what impacts users’ perception of bias on news websites, and to inform the design and development of the subsequent experiments. There were two major obvious potential areas for further research evident from the participants’ responses. First, the impact of the overall characteristics of a website’s design, and second, the impact of the surrounding content on a news website. As the primary focus of this research is to investigate the impact of the design and visual presentation of news online on the perception of bias, it was decided to focus as was originally

intentioned, on the impact of the wider characteristic behind a websites' design rather than the impact of the surrounding content. Consequently, the following studies would focus on the level of professionalism and the visual quality of the presentation and visual aesthetics as they stood out among users' responses.

4.15 Lessons Learned from Experiment One

Several lessons were learned from shortcomings of experiment one which informed the design and development of the second and third experiments and subsequent iterations of the CAFE framework and platform.

- Despite undertaking a priori Statistical Power Analysis test and a post-hoc analysis of achieved power, which revealed a >95% Statistical Power of detecting an effect of 0.42, with an error probability rate of 0.05% which is considered good in the overarching domain of credibility research, the most negative comments of the reviewers focused on the low number of participants. Although 171 completed submissions were recorded, 9 had to be discarded for failing one or more attention questions, while still more had to be discarded due to the then unknown requirement for balanced data to conduct ANOVA. Consequently, future experiments will recruit additional participants to allay reviewers' concerns.
- Having a larger sample population would also likely increase the consistency of results for ascertaining the effects of individual characteristics in the experiment.
- Although no reviewers highlighted it as an issue, having non-applicable or N/A distorted webpage/article combinations in the ANOVA and subsequently in both the initial bias rating Table 9-9 and the re-evaluated bias ratings Table 4-1 meant that it was impossible to properly graph results without either large gaps in the graph or filling in the missing values with averages. These N/As were the result of the individual feature not being present in the original webpage/article combination, thus could not be removed. In future, experiment content will be selected to ensure it is possible to perform all distortions on all webpage/article combinations, thus making it possible to easily graph results.
- Although the participant sample pointed towards a mainly younger student population with high education achievements, the experiment platform did not include the option of "Student" under occupation. This has been remedied for future.
- It has also been recognised that most of the participants were quite young, with an average age of 30.1, with 44.4% of participants between the ages of 18 and 25. While this is representative of the segment of population which traditionally heavily relies on the WWW for their news, future experiments will endeavour to garner the insights of more varied age groups.
- In section 2.2.11 of the State of the Art, the deficiencies of Hong's study are discussed at length. One of the potential issues is that the seriousness of the task, to look for health related information online relating to quitting smoking for a family member or friend, did not suit the aim of the study, to ascertain if structural features of the website's design impact perceived credibility. Judging the

acceptability of such serious information would require users to make judgements of credibility via the central route, the systematic strategy, or the controlled human information process of the ELM, HSM, and the C&AHIP respectively. However, as claimed in these underlying theories, users are likely to rely on visual cues, such as structural elements of a website's design when making judgements via the peripheral route, heuristic strategy, or the automatic human information process of the ELM, HSM, and the C&AHIP respectively. While it is possible to claim that this first experiment suffers the same issue, judging positive or negative bias in climate or weather related news articles is not as serious as judging health related information for a family member, thus making it more likely that participants used the peripheral route, heuristic strategy, or the automatic human information process.

- Another lesson is the possibility that participants in this study had an aversion to Tabloid or Tabloid style news websites. The results in Table 9-28 in the appendix show that participants considered the characteristic of Tabloid, as by far the most likely to increase their perceptions of bias. Results in Table 9-29 also show that participants were more likely to judge the type of content that is often the mainstay of Tabloid newspapers or news websites as being especially biased. It is possible, that given the very high education achievements and other high socioeconomic indicators of the participants in this study, that their aversion is based on a superiority factor and should be treated with a note of caution. This was also a factor in the decision not to pursue this avenue of enquiry in subsequent experiments as it is unlikely that crowdsourcing participants would achieve a more representative sample of the news reading public. This factor will also be considered should this area be explored in future work.
- It is also possible that the impact of some of the distortions was not obvious in themselves. To better demonstrate that the visual presentation of news articles has a significant effect on the perception of bias, more obvious or comprehensive visual treatments should be considered. Consequently, the distortions for subsequent experiments will have a more dramatic effect on larger proportions of the webpage/article combinations being tested.

Recognition of these shortcomings led directly to the lessons learned which influenced the design and development of experiments two and three which are presented in sections 5.2 and section 6.2 respectively.

4.16 Summary of Experiment One Findings

This experiment was designed and developed using an early iteration of the CAFE framework and experiment platform and contributed significantly to its development. The experiment was undertaken to better understand whether individual features of a news website's design, reflected in a news webpage's aesthetic, impact the perception of bias. Research on credibility, of which bias is a core dimension and measure, have shown that individual features can impact perception. Most of the ten MTFS detailed in section 2.2.8 maintain that users can form judgements of credibility using such features via peripheral

routes, heuristic strategies, or automatic human information processes. It is the contention of this thesis that users also form judgements of bias on news websites by such means.

Based on the results of the experiment and subsequent findings, the alternative primary hypothesis was accepted. Users' perception of bias in news articles is affected by their visual presentation. However, when testing the underlying trends supporting the supposition, the results of the secondary hypotheses were less clear. Despite some significant findings and distinct patterns in how users' perception of bias was affected, the overall inconsistency means that it is not possible to categorically reject the any of the three secondary null hypotheses with certainty.

The reflective questions in the experiment also provided a range of insights in an under-researched field. These included the most noticed forms of bias, features of a news article or news homepage which convey or impact perceived bias. The reflective questions also confirmed the original aim of the research, to investigate characteristics behind a website's design, while at the same time revealing a possible future avenue of research, the impact of the type or focus of a news websites content on perceptions of bias.

4.17 Conclusion

Our findings demonstrate that like credibility, users' judgement of bias in a news article online is affected by its visual presentation. And like judgements of credibility, users' judgements of bias may at least partially use the peripheral route, heuristic strategy, or automatic human information process of the ELM, HSM, and C&AHIP, which underlie the ten MTFS of credibility judgement. The reflective questions stage of the experiment also confirmed the original direction of the research, to investigate the characteristics behind a news website's design and webpage aesthetics, specifically, the level of professionalism and the visual quality of the presentation.

“If liberty means anything at all, it means the right to tell people what they do not want to hear.” George Orwell

5 Experiment Two – The Impact of Increasing and Decreasing the Professionalism of News Webpage Aesthetics on the Perception of Bias

5.1 Introduction

As has been demonstrated previously by Fogg, and as part of this research, bias is a core dimension and measure of credibility, especially when users are judging news online (Fogg et al., 2003; Spillane et al., 2018). The State of the Art has shown that the majority of the ten MTFS explaining how users form judgements of credibility online, rely heavily on the peripheral route, heuristic strategy, or automatic human information process from the ELM, HSM, or the C&AHIP when evaluating credibility online (Chaiken, 1980; Petty & Cacioppo, 1986; Shiffrin & Schneider, 1977). Experiment One demonstrated that individual features of a news website’s design, reflected in a news webpage’s aesthetic, can impact perceived bias. Features are individual components of a news webpage that convey information or provide a service, they include advertising, search, login, or integrated social media services. The aim of this experiment is to investigate whether characteristics of a news website’s design, reflected in a news webpage’s aesthetic, also impact the perception of bias. Characteristics are the overarching themes and considerations behind news websites look and feel, which the owners or designers want to convey to users, (see Definitions of Key Terms on page xxix of the preface for more information). Individual characteristics include trustworthy, traditional, bold, active, cutting edge, busy etc. They can be transmitter to the user through coordinated combinations of branding, colour schemes, layout, animation, structure and alignment. To this end, this experiment will investigate the characteristic of Professionalism, which was identified by participants in the reflective questions of Experiment one as likely having an effect on their perception of bias. The importance of design characteristics on judgements of credibility have been noted by many including Sundar who claims they are: *“associated with powerful first impressions of Web site credibility”* (Sundar, 2008).

5.2 Influence from the State of the Art and from Experiment One

In experiment one, as part of a series of reflective questions, participants were asked which common features and characteristics of news websites, homepages and news articles have a particular effect on their perception of bias, see section 4.14.3.3 and section 9.20.4 of the appendix. The thirty-eight items which made up the list for website design characteristic were selected from existing literature on characteristics of successful websites, including those specifically relating to news websites, website satisfiers and

dissatisfies, items relating to trust and credibility, and research relating to the visual identity of commercial and alternative news websites. The four items which stood out most as reducing perception of bias were; Seriousness, Professionalism, High Quality Aesthetics, and Good Design. The first two, seriousness and professionalism are primarily considered as characteristics of the source, or of the organisation behind the website, which is sometimes known as “the sponsor” (K. S. Freeman & Spyridakis, 2004; Metzger, Flanagin, Eyal, Lemus, & Mccann, 2003; Sundar, 2008). The second two, i.e., high quality aesthetics and good design, could be considered surface level characteristics of the means of dissemination as they focus purely on the visual quality of the website. It was decided to investigate professionalism in the second experiment. First, because it has regularly been shown to be a factor of, or to impact perceived credibility (as detailed below), of which bias is a core dimension and measure. Second, its importance has regularly been highlighted in the aforementioned MTFS in user judgements of credibility online such as by (Wathen & Burkell, 2002). As a result, there is a body of literature on the subject. Therefore, it was felt that any findings from this experiment on professionalism would be indicative of the results of a study on seriousness. High Quality Aesthetics and Good Design will be investigated in the third experiment detailed in chapter 6. These two terms are very closely related and focus purely on the visual quality of the website.

In early research in the domain, which focused on the credibility of source and the believability of message, professional or professionalism was consistently seen as an important factor of the construct (Chaiken & Eagly, 1976; Whitehead, 1968). Though it must be noted that it was not included in the item scales and resulting factor analysis of Gaziano and McGrath’s highly cited study (Cecilie Gaziano & McGrath, 1986). Later, professionalism was also seen as an important characteristic in studies of credibility in website design. In Freeman and Spyridakis’s review of website characteristics that affect perceived credibility, they highlight several previous studies which focus on the impact of professionalism (K. S. Freeman & Spyridakis, 2004). Their user study focused on design elements that impacted users’ perception of credibility. Among the 148 comments were several references to “professional design” as positively impacting perceived credibility of a website. An earlier study by Fogg et al., on elements that affect website credibility also collected survey responses such as “The site looks professionally designed” which positively affected perceived credibility (Fogg et al., 2000, 2001). Several studies have also used professionalism as a measure of credibility, including Flanagin and Metzger’s study on website features and Chung’s study on the anatomy of the credibility of online newspapers (Chung et al., 2010; Flanagin & Metzger, 2007). Long and Chiagouris’s study on the credibility of non-profit websites used “The site looks professionally designed” as a measure of credibility (Long & Chiagouris, 2006). A study by McKnight and Kacmar also found that participants perception of credibility during the introductory stage of exploring a website, which was based on the visual quality or professionalism of the website’s design, positively influenced perceptions of credibility when users were further exploring the contents of a website (McKnight & Kacmar, 2007).

Most research into the impact of professionalism on bias, has been on professionalism in journalistic bias. Soloski argues that professionalism among journalists is an effective means of preventing overtly biased or ideologically slanted news (Soloski, 1989). Dennis also argues that professionalism among journalists trumps partisanship: “*Our press is guided more by professionalism than by partisanship*” (Dennis, 1997). In contrast, Baron, while recognising that bias could be constrained by professionalism, takes a much more pessimistic view. He highlights the conflict between demand for high quality news from sceptical readers versus the reduced cost of low-quality biased news. He also notes that journalists may be tempted to bias the news they produce for career advancement, and that while publishers can control this, it may result in less profits, at least in the short term (Baron, 2006).

In contrast, very little work has been undertaken investigating the impact of professionalism on bias in the dissemination of news online. Two exceptions are Muddiman and Stroud’s investigation of professionalism norms in relation to bias in comment sections in news websites, and Orellana-Rodriguez and Keane’s review of the professional reporting of news on Twitter with a focus on bias (Muddiman & Stroud, 2017; Orellana-Rodriguez & Keane, 2018). No work was found by the author investigating the impact of the professional presentation of news online on the perception of bias. This is despite a range of studies demonstrating the impact of the visual appeal, and in particular first impressions. These include studies by Robins and Holmes who demonstrated that websites with increased levels of aesthetic treatment are perceived as more credible than those with low treatment. They also demonstrated that users make judgements of credibility in as little as 3.42 seconds, while Lindegaard et al. have demonstrated that users make judgements about the acceptability of a website as an information resource in as little as 50ms (Lindgaard et al., 2006; Robins & Holmes, 2008).

Despite this, there appears to be no work on the impact of the characteristics of a website’s design or a webpage aesthetic, such as their professionalism, quality of design, dynamism, gaudiness, or how complementary or the suitability of its colour scheme, on the perception of bias or on the overarching domain of credibility.

5.3 Experiment Objectives and Contributions

5.3.1 Objectives

This experiment partially satisfies two of the research objectives. First, it provided continued opportunities to refine the design and development of the CAFE framework and experiment platform as stipulated in:

- **Objective 2:** To design, develop, and test an innovative framework and platform that supports a range of experimental designs to detect and record user’s perception of bias.

Second, the experiment was designed to satisfy the second clause of the third research objective, namely, to investigate whether characteristics of a news website's design, reflected in a webpage's aesthetic, impact perceived bias.

- **Objective 3:** To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impacts perceived bias.

5.3.2 Contribution

Along with experiment one, this experiment forms the first half of the Major Contribution, namely the identification of features and characteristics.

- **Major Contribution:** The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

5.4 Research Ethics Approval

Research ethics approval for this experiment was granted by the Research Ethics Committee of the School of Computer Science and Statistics, Trinity College Dublin.

5.5 Use of Website Titles to Describe Results and Findings

As stated in section 4.5, it should be noted that in this and the two other experiments described in this thesis, news website names and titles are used to describe results, e.g. the BBC was the most or least biased. This was done for simplicity and to differentiate the nine distorted webpage/article combinations. Any results and findings are not reflective of the publication in reality, i.e. if a news website such as Reuters is discovered to be the most or least biased in experiments two or three, that is most likely the result of the news articles, which were deliberately edited to create a range of biased content and were inserted into each webpage, and the distortions applied to them. It is no reflection of the news agency itself. There is a chance some of it can be ascribed to participants recognising the news agency and their positive or negative feelings towards it.

5.6 Experiment Design

The experiment was designed and deployed using an updated version of the CAFE framework and experiment platform that was used in the first experiment. This experiment was set up as a 9x7 within subject incomplete counterbalanced measures design. Incomplete counterbalancing was achieved by arranging the webpages and distortions, D0–D6, in a reduced form Latin square. Participants were

randomly assigned to one of 9 diagonal paths through the Latin square that intersected with 9x7 distorted webpage/article combinations, thus ensuring that each participant experienced each news webpage/article and each distortion once. To reduce carryover effects and reduce the impact of task fatigue, once assigned to a path, the distorted webpage/article combinations the participant would encounter were displayed to the participant in random order. Two attention questions, to determine continued diligence to the task were also added to each participant's path. The independent measures in this study were the webpage/article combinations, and the distortion. Combined these make a (distortion) x (webpage/article) combination. Perceived bias is the dependent measure.

5.6.1 Hypotheses

5.6.1.1 Primary Hypotheses

This experiment was conducted under two primary hypotheses, which are based on the influences from the state of the art presented in section 5.2 which in turn are based on the results of the reflective questions from the last experiment discussed in section 4.14.3.3 of the last chapter and presented in section 9.22 of the appendix. The results of the reflective questions revealed several characteristics which participants believed could impact their perception of bias in a news article. Based on these results it was decided to investigate professionalism of the design and aesthetics. A related, though separate factor quality of the aesthetics is investigated in the third experiment presented in chapter 6. The differences between these and why they were investigated are explained in section 5.2.

The first hypothesis focused on the impact of increasing the professionalism of news webpage's aesthetics on perceived bias.

H₀ Perceived bias will not be decreased due to an increase in the professionalism of the news webpage's aesthetics.

H_A Perceived bias will be decreased due to an increase in the professionalism of the news webpage's aesthetics.

The second, focused on the impact of decreasing the professionalism of the webpage's aesthetics on perceived bias.

H₀ Perceived bias will not be increased due to a decrease in the professionalism of news webpage's aesthetics.

H_A Perceived bias will be increased due to a decrease in the professionalism of news webpage's aesthetics.

It should be noted that the second primary hypothesis is not the opposite of the first. The first is focused on investigating whether increasing the professionalism of the presentation of a news webpage decreases the perception of bias in the news article. If it does not, the null hypothesis is accepted. This means that the bias rating is unchanged, not decreased. The opposite is true for the second hypothesis. Rejecting the alternative and accepting the null means that the bias ratings were not increased when the professionalism in the aesthetics of the webpage was decreased. It is possible that bias can only be affected in one direction due to changes in its presentation. It is also possible that the level of bias in an article will affect the outcome.

5.6.1.2 Secondary Hypotheses

The experiment was also designed to answer two secondary hypotheses. The first aimed to ascertain whether there is a statistical difference in the perception of bias between the most and least professional distorted versions of the same webpage/article combinations.

H₀ There will be no significant difference in the perception of bias between professional and unprofessional versions of the same news website.

H_A There will be a significant difference in the perception of bias between professional and unprofessional versions of the same news website.

This hypothesis was designed to ascertain whether any potential impact of the primary hypotheses is dependent on the level of bias in a news article itself, rather than the webpage.

H₀ The impact of the professionalism of the aesthetics will be inconsistent, regardless of the level of bias in the article.

H_A The impact of the professionalism of the aesthetics will be consistent regardless of the level of bias in the article.

5.7 Methodology

The experiment required participants to rate the amount of bias in nine distorted webpage/article combinations. Three each from traditional print, news magazine, and international multi-format news agencies. Each of the nine news articles were modified to create a range from relatively unbiased to extremely biased. These were then paired with a news webpage for the duration of the experiment. Each of the nine webpage/article combinations was then subject to a series of distortions to progressively and cumulatively reduce the professionalism of the visual aesthetics surrounding each article. Including the control, a 9x7 incomplete counterbalanced measures design was achieved. 405 participants were crowdsourced to partake in the experiment. More information on participant recruitment and filtering can

be found in section 5.7.8. Perceived bias was rated on a VAS from 0 (unbiased) to +100 (extremely biased). The choice of a mono-directional scale over a bi-directional scale is detailed in section 5.7.11.

5.7.1 *Experiment Content*

5.7.1.1 *Websites and Webpages*

A single webpage from each of the nine news websites used in experiment one was used in this experiment. The selection criteria for these websites was detailed in section 4.7.1.1 of the previous chapter. Another reason to use these websites is that their code base was already familiar. To select the individual webpage from each of the chosen websites for inclusion in the experiment, a search was undertaken within each website to find the most recent article relating to the same or a similar topic as the article that would be inserted. Thus, many features of each webpage would be related to the topic of the article, such as section headers, related content links, and breadcrumb trails. Each webpage was paired with a news article for the duration of the experiment. The news article was inserted directly into the HTML of each webpage so that the website's CSS rendered it exactly like one of the standard news articles from each website. Seven distorted versions of each webpage/article combination, D0–D6, were used in the experiment. These were created by applying a series of cumulative distortions to each webpage/article from the nine websites.

- D6: A branded, undistorted version including the website's original name and logo.
- D5: An unbranded, yet otherwise undistorted version of the original news webpage, the most professional distortion.
- D4 to D1: These distortions were incrementally less aesthetically professional versions of D5
- D0: This was the control, a plain text version of each news article and headline with just `<h1>` and `<p>` tags and no other webpage design applied. This established a ground truth bias rating for each article.

At no point during the experiment did the textual content or headline of each article change. As the articles in this experiment have different levels of bias, they were not compared to each other to satisfy the experiment hypotheses, rather they were compared to other versions of themselves. The range of biased articles and the three different categories of websites were used to ascertain if any effect discovered was common across different levels of biased content and types of news website. As a result of this distortion process it was possible to measure the cumulative impact of increasing or decreasing the professionalism of a news webpage's aesthetic treatment.

Like the first experiment, it was decided to use a single webpage from the 9 individual news websites. This decision was based on the need to reduce experiment complexity and the number of participants needed costs to obtain a sufficiently high experiment power. This would also reduce crowdsourcing costs.

5.7.1.2 *Creating Biased News Articles*

To properly test the impact of increasingly professional aesthetics on the perception of bias, it was decided to use a range of news articles from relatively unbiased to extremely biased. The main reason for this was to ascertain whether any effect resulting from reducing or increasing the level of professionalism of the aesthetics on the perception of bias within a news article, was uniform across news articles with different levels of bias. Testing for this is important as news articles have different levels of bias. The different levels of bias within the news articles in the experiment would also not prevent or inhibit the statistical analysis necessary to satisfy the experiment hypotheses. A Two-Way Repeated Measures ANOVA, which was envisioned in the experiment design, includes the statistical test Simple Main Effects for Distortions which focuses on within website comparisons of the same webpage/article combinations with different distortions applied. Of course, if the analysis necessary to satisfy the experiment hypothesis depended on comparisons between websites, then the different levels of bias would affect the outcome, or at the very least require that the level of bias within a news article be treated as a covariate. This approach was not utilised in the first experiment as the distortions were not cumulative, thus it would have been harder to isolate any effect or attribute it to any particular distortion.

Nine news articles were taken from mainstream quality news sources in May 2017. The articles were selected to be un-emotive by avoiding topics such as religion, immigration, vaccination etc., which may have stoked existing strong feelings within participants. In each case the articles were well written and had no known controversies surrounding the topic. Each was shortened to circa 200 words to increase speed of reading and to reduce participant fatigue. This was achieved by truncating the article at the end of a sentence close to the desired length of 200 words. They were then modified to ensure a spread ranging from relatively unbiased to extremely biased. All textual modifications to the news articles and headlines were conducted prior to the experiment beginning. At no point did the text or headlines of any article change during the experiment.

To create the range of articles from relatively unbiased to extremely biased, the articles and their headlines were modified to add the two main classes of bias classified by Recasens et al., epistemological bias and framing bias (Recasens et al., 2013). The authors write that epistemological bias is comprised of factive verbs, entailments, assertive verbs and hedges. Framing bias is comprised of subjective intensifiers and one-sided terms. Consequently, a phrase such as: “*Morocco’s king will...*” became “*Morocco’s visionary king will...*”. In an iterative process, the articles were modified to ensure a range from relatively unbiased to extremely biased. The modified titles of the articles and their designed bias rating can be seen in the first two columns of Table 5-1.

To evaluate this process, a pre-test with 24 participants was conducted. They were tasked with reading each article and rating the perceived bias. The results are shown in the pre-test bias rating column in Table 5-1. The process was successful overall as the results show a strong correlation between the designed level of bias and the pre-test bias ratings.

Only one article, relating to solar energy use in Morocco, was considered more biased by the pre-test participants than it was designed to be. This may be because the article highlighted the role of an individual, in this case the King of Morocco, in the project. It should be noted that while there is some inconsistency, the pre-test experiment was undertaken to ensure that there was a good spread of biased content for the purpose of the final experiment. These results are not used in the final analysis. The D0 control ground truth bias rating was used in the experiment analysis.

The last column in Table 5-1 depicts the ground truth bias rating each article received during the crowdsourced experiment. D0 is a plain text version of each article and headline which had only <h1> and <p> tags. As can be seen from Table 5-1, there is strong correlation in the results of the designed bias and pre-test bias ratings.

Table 5-1 The Article Titles / Headline, Designed Bias Rating, the Pre-Test Bias Rating, the Website each Article was Assigned to, and the D0 Control Ground Truth Bias Rating for each webpage/article combination from the crowdsourced experiment.

Article Title / Headline	Designed Bias Rating	Pre-test Bias Rating	Website Each Article was Assigned to	D0 – Control Ground Truth Bias Rating
EU Commission to allow phone companies increase bills on working class while giving free calls to wealthy	90	89.78	Guardian (Print)	70.51
Morocco to switch on first phase of world's largest solar plant	60	74.56	Economist (News Magazine)	30.27
Poison air in holy city of Varanasi 'most toxic in India'	80	71.33	Al Jazeera (News Agency)	42.77
Bangkok cleans up its act	70	61.56	Telegraph (Print)	54.27
New Zealand Prime Minister John Key resigns in shock announcement at weekly news conference	50	59.22	The Spectator (News Magazine)	36.02
Progress on malaria deaths at serious risk without big boost in funding, UN warns	40	42.11	BBC (News Agency)	31.87
Pakistan fights devastating malnutrition with mass food-fortifying programme	30	39.78	Independent (Print)	26.98
U.S. import prices post largest drop in nine months worrying economists	20	31.00	New Statesman (News Magazine)	20.02
Two-wheel takeover: bikes outnumber cars for the first time in Copenhagen	10	22.22	Reuters (News Agency)	22.36

The designed and pre-test bias ratings did not however correlate strongly with the D0 bias rating from the experiment. However, as the aim was to have a range of biased content in the crowdsourced experiment, the article distortion process was successful. The three forms of bias rating, designed, pre-test, and ground truth are also depicted in Figure 5-1 along with the webpage they were assigned to. Figure 5-1 is an alternative depiction of the information shown in Table 5-1.

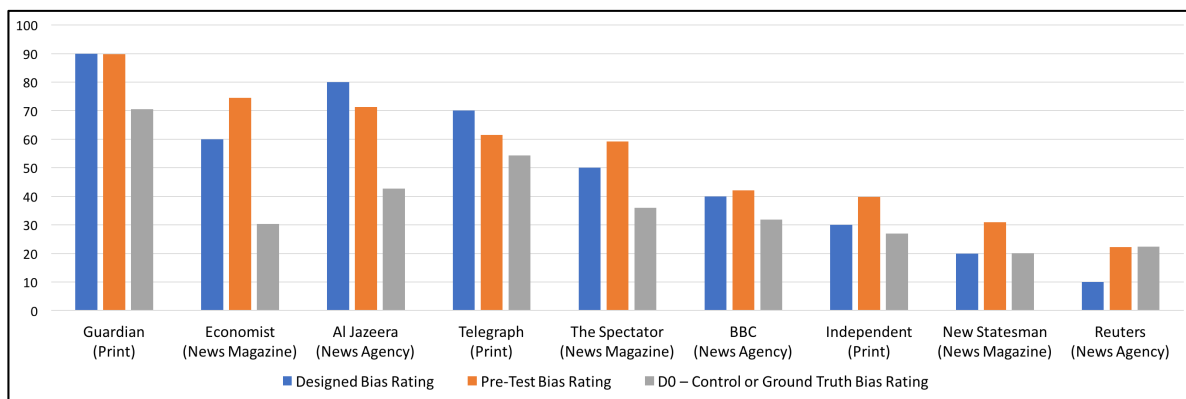


Figure 5-1 The Designed Bias Rating, Pre-Test Bias Rating, and D0 Ground Truth Bias Rating from Table 5-1.

Unlike the previous experiment, it was decided to use a range of biased content. This was based on the desire to see if the effect of the level of professionalism was consistent. This would make it more applicable to news website designers and online editors.

5.7.1.3 Pairing of News Webpages and News Articles

To investigate the impact of different levels of biased articles on different categories of websites, print, news magazine, and international multi-format news agencies, a high, medium, and low-level biased article was assigned to each category. The article would only be displayed in its assigned webpage for the duration of the experiment creating a webpage/article combination, which was then subject to distortions. The websites to which each article was assigned to can be seen in Table 5-1.

Each of the three categories of news website (traditional print, news magazine, and international multi-format news agency) was assigned an article with high, medium, and low levels of bias which was determined from the pre-test results, shown in the middle column in Table 5-1 and shown in orange in Figure 5-1. Thus, as can be seen in Table 5-1 and Figure 5-1, the three most biased articles (by pre-test bias ratings) were assigned one each to a traditional print, news magazine, and international multi-format news agency. This process was repeated for the three mid-level biased articles and for the three least biased articles.

Although the primary focus was on investigating the impact of the level of professionalism of the design and aesthetic, it was decided to utilise edited real news articles paired with distorted news webpages to create a realistic browsing experience. An alternative approach could have been to blank out the content or simply replace it with placeholder content to focus the participant on the aesthetic created by the webpage. However, it was felt that this would not be representative of a real browsing experience and therefore the results would be less useful to news website designers and online editors.

5.7.1.4 Images in News Articles

Most of the news articles were accompanied by an image which was placed at the top of each article. However, in some instances, an article being used in the experiment did not have an image. To ensure consistency in the experiment content, an appropriate image therefore had to be added to any news article without one. It was important that these images did not impact the experiment. To identify suitable images, a search was first undertaken for images related to the topic of the article, in the same website the article came from. Once 3-5 images for each of the missing articles were identified, the following selection criteria was used to select the final image. 1) It specifically related to the topic of the article. 2) It did not depict content which might impact the perception of bias in the article. 3) It fit the tone of the article. Following this, a picture of a port and cranes was selected for the article on US imports, and a picture of a cyclist was selected for the article on increased use of bikes in Copenhagen. An alternative approach would have been to remove all images. However, this was discounted based on the fact that many news articles include images and it was worthwhile including them.

5.7.2 Professional Website Design and Webpage Aesthetics

The aesthetic of an individual news webpage is primarily derived from the website's design, though a number of other factors such as third-party advertising, social media, the content being displayed, and personalisation all play a role. Even the viewers' setup, including their computer screen, browser, resolution, and browser extensions such as ad-blockers can all impact a webpage's aesthetics. This study was not designed to measure the professionalism of the website's design, but rather the professionalism of the visual aesthetics of a single webpage. However, the close relationship between a website's design and a webpage's aesthetics must be acknowledged. As such, this chapter primarily uses the term 'professional webpage aesthetics', however it occasionally uses the term 'professional website design' when referring to the overarching influence.

Currently there is no theoretically grounded benchmark or means of measurement for professional website design or professional webpage aesthetics in the literature, although the former term is used within credibility research. In a study into how users appraise health information online, professional design was shown to be a significant concern of information seekers (Eysenbach & Köhler, 2002). Ladhari and Michaud maintain that professional website design is at least partially responsible for positive attitudes towards a website (Ladhari & Michaud, 2015). Stephens, in his framework for identifying website elements that positively impact trust, a closely related research area, also highlights the importance of professional design (Stephens, 2004). Section 5.2 of this chapter also highlighted several studies which demonstrate the importance of 'professional design' to the users judging credibility online.

Design and aesthetic are also considered core, though often interchangeable dimensions of 'website quality' which is used to measure the performance of websites. Website quality is a multidimensional

construct, with the number of dimensions varying between three (Rocha, 2012) and twelve (S. Kim & Stoel, 2004a). Though some, such as Lin (H.-F. Lin, 2007) use the term ‘design’, it may also be known as visual appeal, appearance, aesthetics, visual presentation, visual design, or graphic design etc. Loiacono is one such who uses the term visual appeal to describe: “*the aesthetics of a web site*” (Loiacono et al., 2002).

While it is difficult to describe or quantify what exactly constitutes professional aesthetics, it is relatively easy to describe unprofessional news website aesthetics. Issues with alignment, broken code, low quality images, bad or conflicting colour schemes, gaudiness, mismatched fonts or text attributes, haphazard positioning, and low quality or gaudy advertising etc. all contribute to a sense of unprofessional news website design and webpage aesthetics. News organisations cannot afford to have their websites judged ‘unprofessional’ and therefore spend millions continuously improving their online presence to maintain and increase their audience. The news websites used in this experiment are among the most popular in the world. The Guardian alone reported 167 million monthly unique browsers in June 2016 (GNM press office, 2016). It was therefore judged that original, undistorted webpages from each of these websites would constitute aesthetically professional.

5.7.3 *Distortions*

The aim of the distortion process was to progressively, and as much as possible, commonly degrade the level of professionalism in defined stages (D4 - D1) so that each could be examined in the experiment. To begin, a review of the websites to be used in this experiment, and others including local and national news websites was conducted to establish the differing aspects of professionalism conveyed to the user. The focus was on the medium and not on the content of each website. Therefore, incendiary or bellicose headlines, the message in the article text, or the style of images, such as tabloid style images of celebrities, was ignored. Table 5-2 contains the outcome of that review, including the means of affordance by which the aspect of professionalism is conveyed to the user and the action necessary to negatively impact it. Aspects such as reputation were not included as they relate to the organisation rather than the means of dissemination.

A two-pronged approach was used to degrade each webpage. First, technical support for the underlying technologies, which Sundar claims are the means of providing such affordances (Sundar, 2008), was removed by deleting the underlying code. Second, every image used in each webpage, including in article photographs and vector images used in each webpage, were degraded by reducing file size, bit rate, and DPI. A screenshot was then taken of this degraded page, which was then again degraded by reducing its file size, bit rate, and DPI.

Table 5-2 Depicts the seven main factors of professional news website design resulting from the review, the main underlying design or technical means of conveying it, and the action required in the distortion.

ID	Aspect of (Un)Professionalism Conveyed to the User	Means of Affordance	Action to Reduce Professionalism
A1	Overall design / colour scheme	Design, layout, images, colour schemes	Add issues with alignment and layout. Reduce the quality of the colour scheme
A2	Technical quality of the website	HTML/CSS/ Java Script	Reduce or remove elements that convey technical ability
A3	Design appropriateness, reserved, custom, unique or non-standard features or elements	Image Quality, fonts, interactive features and widgets	Degrade quality of images, introduce basic or standard elements e.g. fonts or third-party services
A4	Money and/or technical ability in the organisation behind the website	Technical ability	Introduce obvious technical issues
A5	Credibility, focus on the news article	Size, prominence, quality of advertising	Increase amount and prominence of advertising
A6	Excessive calls to action on the user	Sharing, signup, user interaction or other methods of connecting	Increase the amount or size of calls to action
A7	Expertise's on the topic, history dealing with the subject	Information scent via links to supporting content	Remove or reduce supporting links

A description of each distortion can be seen in Table 5-3. Figure 5-2 and Figure 5-3 depict the distortion process as applied to the Al Jazeera webpage/article combination. This distortion process was designed to reduce the professionalism of the aesthetics of each webpage/article combination. It was also designed to be replicable by focusing on removing underlying technologies or by reducing the quality of images and graphics. While the colour schemes of the news websites were raised several times during the review as a means by which professionalism could be conveyed to the user, see Table 5-2, no distortion was applied to this aspect of the websites due to any reductions in professional quality being subjective and difficult to replicate. It should also be noted that not all the distortions had the same level of effect on all websites. How each website was constructed, including which underlying technologies were utilised, impacted the severity of the distortion. However, this could not be helped, and it contributed to the decision to use nine websites in the experiment to provide a larger sample, thus ironing out inconsistencies of effect.

Overall the distortion process was a success. As underlying code supporting the different technologies was removed, issues with alignment, positioning, overlap, and even failure began to appear as is evident in Figure 5-2 and Figure 5-3. In this instance, the D5 distortion on the top left is the most professional of the unbranded versions of the Al Jazeera webpage/article combination. Partially visible is the article it was paired with from Table 5-1 which was designed to be the second most biased article in the experiment. While each of the elements on the left are aligned correctly, creating a high-quality aesthetic look, the elements on the right have begun to move out of line, gaps and spaces have begun to appear, and in some cases the element failed to render in the browser.

Table 5-3 The stages of the cumulative distortion process to reduce the professionalism of the webpage/article combinations.

Stage	Distortion	Distortion Description
1	D6 Original, Undistorted Including Name and Branding	This is an undistorted original version of each webpage/article combination, which includes its original name and branding with only minor distortion to remove comment facilities. Visible comments in comment facilities were removed so that the comment facility was still present, but the opinions expressed in it would not impact the experiment A static image of webpage/article combination was then created
2	D5 Original, Distorted to Remove Name and Branding	Starting with the previous stage, the logo was removed and replaced with generic NewsCom moniker with a circle logo in the primary colour scheme. Also, removed all incidences of the news organisations name from the rest of the webpage, footer, self-promotional advertising etc. A static image of webpage/article combination was then created
3	D4	Starting with the previous stage, remove CSS 3 and CSS 4 support. This was undertaken by finding each tag introduced in CSS 3 and CSS 4 and removing it from the code Reduce the file size, bit rate, and DPI of individual images used in the webpage/article combination Create static image of webpage/article combination Reduce the file size, bit rate, and DPI of the final image A static image of webpage/article combination was then created Reduce the file size, bit rate, and DPI of the final image
4	D3	Starting with the previous stage, remove all elements introduced in HTML 5.1 standard Further reduce file size, bit rate, and DPI of individual images used in the webpage. Resize images to show stretching Remove all custom fonts and replace with Times New Roman Remove all JavaScript support and functionality A static image of webpage/article combination was then created Reduce the file size, bit rate, and DPI of the final image
5	D2	Starting with the previous stage, remove all elements introduced in CSS 2 / 2.1 Further reduce the file size, bit rate, and DPI of individual images. Also, further reduce image size to increase stretching A static image of webpage/article combination was then created Reduce the file size, bit rate, and DPI of the final image
6	D1	Starting with the previous stage, unlink all CSS files but keep HTML elements Further reduce file size, bit rate, and DPI of individual image quality and size A static image of webpage/article combination was then created Reduce the file size, bit rate, and DPI of the final image
7	D0 The Control to establish a ground truth bias rating	Remove all CSS / HTML except for the plain text, <p> and <h1> tags in the article Remove all images A static image of webpage/article combination was then created



Figure 5-2 The D6, D5, D4, D3, D2, D1, and D0 distorted webpage/article combinations for Al Jazeera.

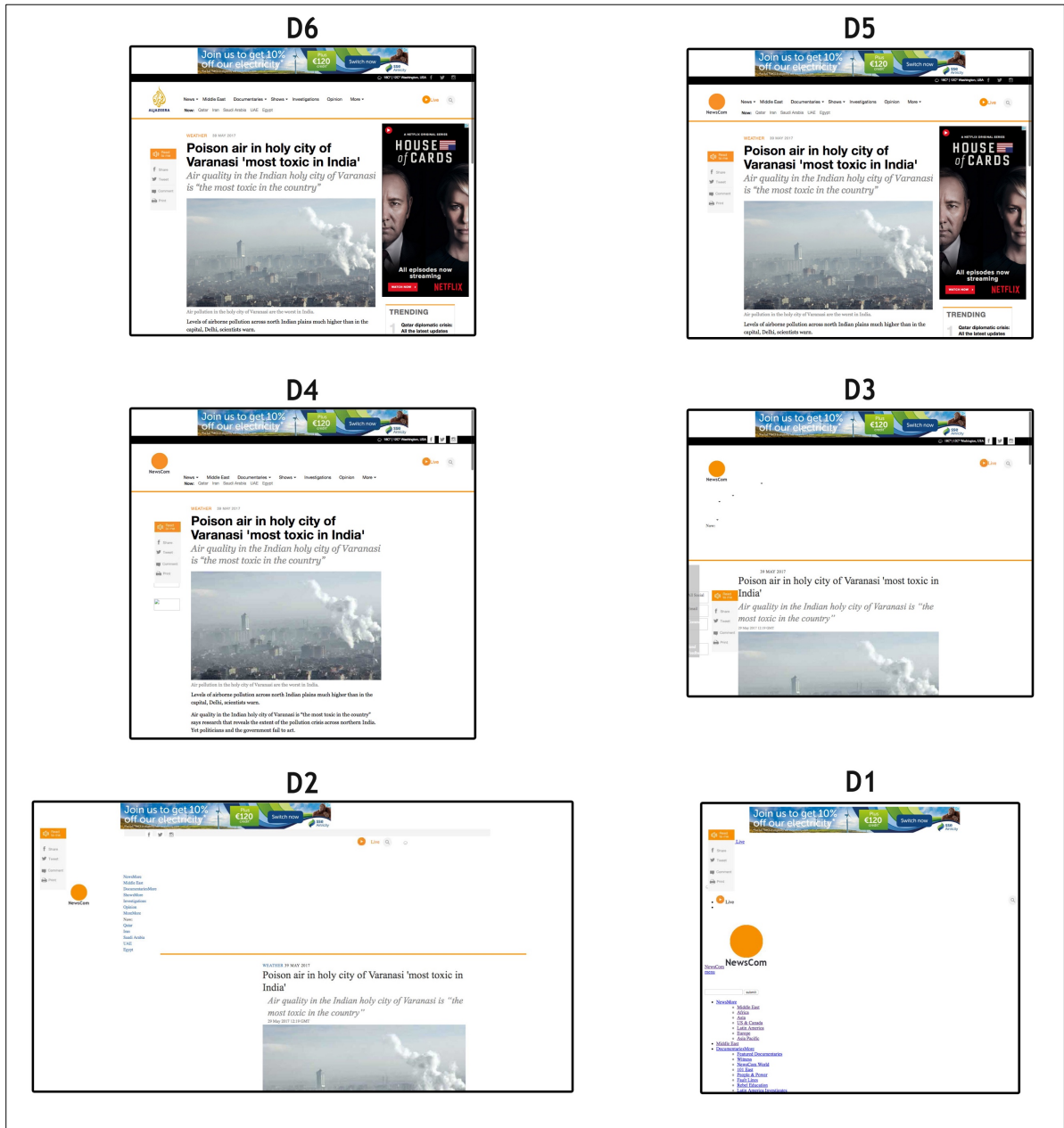


Figure 5-3 The D6, D5, D4, D3, D2, and D1 distorted webpage/article combinations for Al Jazeera.

5.7.4 Measuring the Success of the Distortion Process

To ensure that the distorted webpage/article combinations for each website reflected a range from professional to unprofessional, a pre-experiment evaluation was undertaken. 12 participants were presented with the D1-D5 distorted air webpage/article combinations from each website in random order. They were then tasked with organising them from most to least professional (D5 – D1). The D0 distortions were not included as they were a plain text version of the news article with no design applied, while the D6 distortions were not included as it was just a branded version of the D5 with user comments on the articles included.

Cronbach’s Alpha is a statistical test to measure internal consistency, also known as reliability (Cronbach, 1951). A score of 1.0 indicates 100% consistency. Scores $\geq .9$ are considered excellent, scores $\geq .8$ are considered good, scores $\geq .7$ acceptable, scores $\geq .6$ questionable, scores $\geq .5$ poor, while scores $< .5$ are considered unacceptable.

Table 5-4 depicts the Cronbach’s Alpha scores for the webpage/article combinations. The high Cronbach’s Alpha scores show that the majority of the pre-test participants successfully arranged the distorted webpage/article combinations for each website in the correct order, from most to least professional (D5 to D1). This means that in the majority of cases, for each of the websites, the participants were able to recognise the decreasing level of professionalism in the distorted webpage/article combinations, as they were almost always able to arrange them in or close to the right order. Consequently, it is possible to say that this distortion process was successful in creating a range from professional to unprofessional and vice versa of distorted webpage/article combinations.

Table 5-4 Cronbach’s Alpha Scores for each webpage/article combination. Pre-test participants were tasked with arranging them from D5 to D1.

Article Title / Headline	Website	Cronbach’s Alpha
EU Commission to allow phone companies increase bills on working class while giving free calls to wealthy	Guardian	.949
Morocco to switch on first phase of world's largest solar plant	Economist	.892
Poison air in holy city of Varanasi 'most toxic in India'	Al Jazeera	.960
Bangkok cleans up its act	Telegraph	.970
New Zealand Prime Minister John Key resigns in shock announcement at weekly news conference	Spectator	.951
Progress on malaria deaths at serious risk without big boost in funding, UN warns	BBC	.892
Pakistan fights devastating malnutrition with mass food-fortifying programme	Independent	.897
U.S. import prices post largest drop in nine months worrying economists	New Statesman	.846
Two-wheel takeover: bikes outnumber cars for the first time in Copenhagen	Reuters	.977

5.7.4.1 Website Logos and Branding

In this experiment, it was decided to remove the logos and branding from each webpage and replace them with the generic moniker NewsCom with a circle graphic using the same colour scheme as each webpage. All other branding and incidences of the original news agencies name or logo were similarly removed and replaced. This was done in Stage 2 of the distortion process shown in Table 5-3. As this distortion process was cumulative, Stage 3 building on Stage 2 etc., there was no branding or logos on any of the Distortions D5 – D0 for any of the websites.

5.7.5 *The Experiment Platform*

This experiment was conducted using an updated version of the experiment platform from the first experiment. This was the predecessor to the opensource experiment platform released with the CAFE framework, which is detailed in section 3.5 of this thesis.

5.7.6 *Definition of Bias and Participant Instructions*

The instructions and the definition of bias were designed to direct the participant to the news articles contained in each webpage to assess the amount of bias in it. This was reiterated multiple times to ensure that the focus of each participant was on the news articles. Therefore, any change in the bias rating that the distorted webpage/article combinations received could be attributed to the professionalism of the aesthetics.

To begin with, the advertisement on the crowdsourcing platform stated: *“In this study, you will be asked to rate, from 0 to 100, how biased different online news articles are. 0 being unbiased and 100 being extremely biased.”* Participants were then given the following definition of bias: *“Bias is defined as **‘Deliberate or accidental slant by the journalist, editor or publication to distort reality.’**”* Participants were again presented with the above definition in a popup dialog during the first Instruction Task, and again after they answered the user profile questions. On the second Instruction Task page, another dialog box appeared with the message: *“Some of the webpages may appear broken. However, each contains a News article which you will find by scrolling up or down the page”*. This was to ensure that despite any appearance that the webpage was broken due to reduced professionalism, participants understood that each webpage did contain article and that it was the focus of their task.

This definition was designed to focus the user on the article without specifying the type of news bias they might be exposed to. These include, but are not limited to, presentation, source, agenda setting and/or framing bias. It was also chosen as many of the definitions in the literature focus on the amount and quality of coverage a topic received over time, such as that by Stevenson et al. who conceptualised bias as: *“imbalance or inequality of coverage”* (Stevenson et al., 1973), or too specific, such as Waldman and Devitt’s: *“any systematic slant favoring one candidate or ideology over another”* (Waldman & Devitt, 1998). While some within the literature might disagree with using the term ‘slant’ in a definition of bias, it should be noted that Entman connects the two related but distinct terms and Waldman and Devitt, and Gentzkow and Shapiro use the term slant in their definitions (Entman, 2007; Gentzkow & Shapiro, 2010; Waldman & Devitt, 1998).

5.7.7 *Five Stages of the Experiment*

Like experiment one, this also had five stages including; setup (terms and conditions and participant instructions), profile questions, the bias rating tasks, reflective questions, and submit and debrief. These stages were previously detailed in section 4.7.5 of the last chapter.

5.7.8 *Crowdsourcing*

In total 508 completed submissions were recorded in the database. 47 were rejected for failing one or both attention questions. 32 were rejected for having none or wrong completion codes and/or for completing the experiment within 1-2 minutes, indicating automated bots at work. 10 were not included in the final analysis due to the participants rating >4 webpage/distortion combinations as having 0 bias. Finally, 14 were not included in the analysis due to having missing data, probably due to JavaScript not being enabled on the participant's browser, or to balance the dataset. Therefore, the final dataset consisted of 405 submissions, 45 participants per path, and 45 participant ratings for each distorted webpage/article combination. To ensure balance for the subsequent ANOVA, participants were assigned to one of 9 paths at random by the experiment application. When a path was full it closed automatically.

Participants were recruited through Prolific Academic and paid £1.25 (Palan & Schitter, 2018). Participants were over 18 with English as their first language. Participation was limited to the US to increase the homogeneity of the data and because the suppositions made about the experiment were based on a US population. One of the lessons from experiment one was that too many of the participants were too young, possibly slanting the results. Consequently, experiment two was run in three iterations, each limited to different age groups (18 – 30, 30 – 40, and 40+) over four days in late August 2017.

5.7.9 *Steps to Ensure Validity of Results*

Like experiment one, six actions were taken to increase validity. These were detailed in section 4.7.7 of the last chapter and in the CAFE framework which can be seen in section 9.21 of the appendix.

5.7.10 *Statistical Power and Effect Size*

A Statistical Power Analysis test for ANOVA: Fixed Effects, Special Main Effects And Interactions, was conducted a priori using G*Power (Faul et al., 2007, 2009). To achieve $>.95$ actual power of detecting an Effect Size of 0.25 a sample size of 372 would be required. As learned from the first experiment, ANOVAs required balanced data. Consequently, a multiple of 9 greater than 372 would be required as there were nine paths in the experiment. A post-hoc test to compute achieved power with the 405 participant submissions included in the final dataset revealed a $>.95$ Actual Power to detect an Effect Size of 0.24. Though there is little to compare it to in bias research, this would be considered very strong in the overarching domain of credibility.

5.7.11 Bias Rating Scale - VAS

A VAS, ranging from 0 to 100, was used to measure bias, see Figure 5-4. This type of scale had performed well in the first experiment and consequently was adopted again. However, the first experiment was exploratory in nature and designed to see if a news webpage’s design impacted perceived bias. Thus, a bi-directional scale (-100 negative to +100 positive), paired with a broad definition of bias was used in that experiment. In this experiment, the news articles were modified and pretested to ensure a range of biased content, thus this experiment was not exploratory. Instead, the focus was on identifying to what degree the different levels of distortion impacted perceived bias on a range of biased content. For this reason, a mono-directional VAS was used in this, the second experiment. It was also used in the third experiment for the same reason. The mono-directional VAS shown in Figure 5-4 was placed below each distorted webpage/article combination. Above the scale was the instruction: “Please rate how Biased this News Article is. 0 being unbiased and 100 being extremely Biased”. Each end of the scale was anchored with 0 and 100 with the headings “Unbiased” and “Biased” respectively.



Figure 5-4 Mono-directional Visual Analogue Scale (VAS) from 0 - 100 used in this experiment.

5.8 Participant Profile

405 submissions were evaluated. 51.4% Male, 47.2% Female, 0.7% Other, and 0.7% Prefer not to say. Mean age was 36.48 years old. 34% of participants were <30, validating the earlier decision to run the experiment in multiple iterations to gather a more diverse age profile. Participants were spread across the US with 44 states and the District of Columbia represented. Further details of the sample profile are provided in section 9.24 of the appendix.

5.9 Statistical Analysis and Results – Re-Evaluated Bias Rating

The experiment was set up and deployed on an updated version of the CAFE experiment platform from experiment one. The experiment matrix was a Latin square in reduced form with the two independent measures (webpage/article combinations) and (distortions) arranged on each axis. Participants were assigned at random to a diagonal path through the Latin square, thus ensuring that they experienced no

webpage/article combination or distortion more than once. Each distorted webpage/article combination was rated by 45 participants.

5.9.1 Re-Evaluated Bias Ratings

Like the previous experiment, participants were given the opportunity to comparatively re-evaluate their bias ratings after they had rated three randomly selected webpage/article combinations from the Latin square experiment matrix. The addition of this re-evaluation step was to ensure that participants were happy with their perceived bias rating. The main supposition behind this was that the influence of the visual presentation, such as the professionalism of the webpage's aesthetics, may only become apparent during a side by side comparison with other webpages. This step would also increase the validity of the results as they better reflect the participant's opinion.

In the previous experiment the initial and re-evaluated bias ratings were analysed and presented separately. While this demonstrated the overall consistency of results it failed to discover any patterns in the re-evaluation. The final experiment findings were also derived from the re-evaluated results. As this step worked well to increase validity it was included in the second experiment. However, in this experiment only the final re-evaluated results are analysed to derive the experiment findings.

5.9.2 Main Effects

A 9x7, Two-Way Repeated Measures ANOVA using the Bonferroni correction for multiple comparisons was undertaken using SPSS 24. Pre-test bias scores were not used as covariates for two reasons. First, because they were only recorded to ensure a range of news articles from relatively unbiased to extremely biased, and second, because the focus is on within-website comparisons. That is, each distorted webpage/article combination is only compared to other versions of itself and not to other distorted webpage/article combinations from other websites. An analysis of the studentized residuals for data points ± 3 standard deviations revealed no outliers. A Shapiro-Wilk test ($p > .05$) for normality on the studentized residuals revealed ~40% data was not normally distributed. However, the Shapiro-Wilk test is considered especially sensitive to deviations from normality on data sets > 50 participants. As this dataset has 405 participants, a visual inspection of QQ plots was undertaken which showed that, of the ~40%, half was not normally distributed. Therefore, it was decided to use the ANOVA because it is considered fairly robust to deviations from normality. Mauchly's Test of Sphericity (MTS) ($p > .05$) indicated that the assumption of sphericity had been violated for the two-way interaction. Consequently, the Greenhouse-Geisser (GG) correction is reported rather than the Huynh-Feldt (HF) as the epsilon (ϵ) was < 0.75 .

As expected, the Two-Way Repeated Measures ANOVA revealed no statistically significant two-way interaction between webpages and distortions, $F(23.22, 1021.79) = .861$, GG $p = .655$. This is because the

experiment was set up to detect the cumulative impact of the distortions. It was not expected that a significant two-way interaction would be detected.

The main effects of webpage/article combinations showed a statistically significant difference $F(6.60, 290.28) = 80.38$, $GG\ p = <.0005$. This was a likely due to the differing levels of biased content inserted into each website. However, as the focus is on investigating the impact of distortions, further analysis in the form of Simple Main Effects for Webpage/Article Combinations were not undertaken.

An analysis of the main effects of distortions showed that there was a statistically significant difference $F(5.46, 240.42) = 12.24$, $GG\ p = <.0005$. These are further detailed in the section 5.9.3.

5.9.3 *Simple Main Effects for Distortions*

While Simple Main Effects would usually not be explored without a significant two-way interaction, the focus of this research is on comparing the distorted webpage/article combinations to each other and their respective controls. In SPSS, Simple Main Effects are undertaken by performing multiple One-Way Repeated Measures ANOVAs. Alternatively, one could perform multiple paired samples t-tests, however the increased validity of the One-Way Repeated Measures ANOVA due to the option of the Bonferroni correction is preferred. The assumptions required for Two-Way Repeated Measures ANOVAs are also valid for One-Way Repeated Measures ANOVAs. It should also be noted that due to the central limit theorem, One-Way Repeated Measures ANOVAs can still provide valid results even when the distribution of the data is very non-normal. Table 5-5. shows the results of One-Way Repeated Measures ANOVAs – Simple Main Effects for Distortions, using the Bonferroni correction for multiple comparisons for each of the websites.

The Simple Main Effects for Distortions revealed four significant results. Telegraph (D1 – D6), Independent (D1 – D6 and D0 – D1) and Spectator (D0 –D1).

5.9.3.1 *Guardian*

MTS ($p >.05$) indicated that the assumption of sphericity was met for the two-way interaction, $\chi^2(2) = 21.28$, $p = .382$. Mean bias rating for distortions was significantly different for the Guardian, $F(6, 264) = 2.352$, $GG\ p = .031$, $\text{partial } \eta^2 = .051$. While a significant value of .031 was reported a closer analysis of Pairwise Comparisons revealed no statistically significant results. This is likely due to the different sensitivity of the two tests. Consequently, it was decided to err on the side of caution and not report this as a significant result.

5.9.3.2 *Telegraph*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 9.632$, $p = .974$. Mean bias rating for distortions was significantly different for the Telegraph, $F(6, 264) = 4.294$, GG $p = < .0005$, partial $\eta^2 = .089$. An analysis of the pairwise comparisons showed a decrease in perceived bias from D1, 59.44 ± 25.07 to D6, 41.00 ± 26.70 (95% CI 1.71 to 35.17), GG $p = .019$.

5.9.3.3 *Independent*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 26.595$, $p = .148$. Mean bias rating for distortions was significantly different for the Independent, $F(6, 264) = 3.328$, GG $p = .004$, partial $\eta^2 = .070$. An analysis of the pairwise comparisons for the Independent showed a decrease in perceived bias from D1 44.62 ± 24.96 to D0 26.98 ± 24.16 (95% CI 2.66 to 32.62), GG $p = .009$. The pairwise analysis also showed a second significant reduction in perceived bias from D1 44.62 ± 24.96 to D6 30.40 ± 25.27 (95% CI 0.85 to 32.70), GG $p = .031$.

5.9.3.4 *Economist*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 18.912$, $p = .529$. Mean bias rating for distortions was not significantly different, $F(6, 264) = 1.480$, GG $p = .185$, partial $\eta^2 = .033$.

5.9.3.5 *Spectator*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 13.221$, $p = .868$. Mean bias rating for distortions was not significantly different for the Spectator, $F(6, 264) = 1.725$, GG $p = .115$, partial $\eta^2 = .038$. While no significant difference was revealed in the Simple Main Effects for Distortions for the Spectator, an analysis of the pairwise comparisons showed a significant decrease in perceived bias from D1 53.04 ± 28.88 to D0 36.02 ± 26.99 (95% CI 0.054 to 33.991), GG $p = .049$. This is likely due to the different sensitivity of the two tests.

5.9.3.6 *New Statesman*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 26.197$, $p = .160$. Mean bias rating for distortions was not significantly different, $F(6, 264) = 2.061$, $p = .058$, partial $\eta^2 = .045$.

5.9.3.7 *Al Jazeera*

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 20.329$, $p = .439$. Mean bias rating for distortions was not significantly different $F(6, 264) = 1.342$, $p = .239$, partial $\eta^2 = .030$.

5.9.3.8 BBC

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 8.939, p = .989$. Mean bias rating for distortions was not significantly different, $F(6, 264) = 1.798, p = .100$, partial $\eta^2 = .039$.

5.9.3.9 Reuters

MTS ($p > .05$) indicated that the assumption of sphericity was met for the one-way interaction, $\chi^2(2) = 15.390, p = .755$. Mean bias rating for distortions was not significantly different, $F(6, 264) = .913, p = .486$, partial $\eta^2 = .020$.

5.10 Results

The significant results of the statistical analysis detailed in the previous section and shown in Table 5-5, Figure 5-5, and Figure 5-6 demonstrate that the professionalism of a news webpage's aesthetic impacts perceived bias within the news article they contain. Table 5-5 depicts the results of One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions with the Bonferroni correction for multiple comparisons. The intersecting cells (webpage/article X distortion) show the Mean bias score and Standard Error. This Mean bias score is the Mean of the 45 participants bias rating for each webpage/article distortion.

- **D0:** These are the controls, plain text versions of each article to establish a ground truth bias rating. Note the Mean bias ratings in the D0 column are the same as those shown in the right most column of Table 5-1.
- **D1 – D4:** The D1s are the most distorted and least professional, while the D4s are the least distorted, more professional.
- **D5:** These are copies of the D6s distorted to remove name and branding.
- **D6:** These are original undistorted versions of each webpage/article combination with all logos and branding.

Were significant differences were found between two distorted webpage/article combinations, the F Statistic and Pairwise Significance p value are shown in the cells beneath. These are also highlighted in colours (shown below) corresponding to their respective colours in the graphs in Figure 5-5 and Figure 5-6. Four significant results were found.

- | | | |
|-----------------------|----------------------------|---------------|
| • Telegraph: | D1 – D6 | Orange |
| • Independent: | D0 – D1 and D1 – D6 | Grey |
| • Spectator: | D0 – D1 | Blue |

Table 5-5 Results of One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions with the Bonferroni correction for multiple comparisons.

	D0 Ground Truth	D1	D2	D3	D4	D5 Original, Distorted to Remove Name and Branding	D6 Original, Undistorted Including Name and Branding
Guardian	M 70.51 SE 4.33	M 76.91 SE 2.86	M 74.04 SE 3.20	M 74.47 SE 3.18	M 63.13 SE 4.72	M 63.29 SE 4.14	M 64.96 SE 4.08
Telegraph	M 54.27 SE 4.16	M 59.44 SE 3.74	M 59.53 SE 3.59	M 58.98 SE 4.15	M 43.18 SE 4.22	M 44.60 SE 4.48	M 41.00 SE 3.98
		F 4.294 p .019					F 4.294 p .019
Independent	M 26.98 SE 3.60	M 44.62 SE 3.72	M 35.33 SE 3.59	M 38.84 SE 3.87	M 30.04 SE 3.48	M 30.40 SE 4.10	M 27.84 SE 3.77
	F 3.328 p .009	F 3.328 p .009					
		F 3.328 p .031					F 3.328 p .031
Economist	M 30.27 SE 4.10	M 46.04 SE 4.43	M 43.09 SE 4.54	M 41.84 SE 4.19	M 42.27 SE 4.88	M 41.00 SE 4.01	M 36.18 SE 4.19
Spectator	M 36.02 SE 4.02	M 53.04 SE 4.31	M 50.33 SE 3.99	M 46.35 SE 4.48	M 48.22 SE 4.28	M 41.80 SE 4.64	M 45.73 SE 4.33
	F 1.725 p .049	F 1.725 p .049					
New Statesman	M 20.02 SE 3.79	M 31.73 SE 4.32	M 31.11 SE 4.33	M 28.00 SE 4.41	M 32.35 SE 3.78	M 19.47 SE 2.67	M 23.56 SE 3.68
Al Jazeera	M 42.77 SE 4.56	M 54.20 SE 4.47	M 52.09 SE 3.66	M 50.58 SE 4.62	M 40.51 SE 4.42	M 47.87 SE 4.26	M 45.29 SE 4.12
BBC	M 31.87 SE 4.27	M 41.51 SE 4.12	M 37.09 SE 4.11	M 37.13 SE 3.39	M 41.87 SE 3.68	M 26.53 SE 9.92	M 32.78 SE 4.58
Reuters	M 22.36 SE 3.73	M 33.33 SE 4.13	M 25.69 SE 3.85	M 25.58 SE 4.20	M 24.07 SE 3.36	M 23.49 SE 3.77	M 22.53 SE 3.96

It should be noted, that each of the news articles was modified to create a range of biased content from relatively unbiased to extremely biased. This is described in detail in section 5.7.1.2 and in section 5.7.1.3.

Figure 5-5 is a graph of all the Mean bias scores shown in Table 5-5. Perceived Bias (0 to 100) is on the Y Axis. The distortions, D0 - D6, are on the X axis. D0 is the control, a plain text version. D1 is the most severe distortion and least professional, D2, D3 and D4 are less severe and increasingly professional. D5 is an unbranded version of each webpage while D6 is a branded version. Clearly visible is a curve showing an increase in perceived bias. Most evident is the marked increase from the D0 plain text version of each webpage/article distortion to the D1, least professional version. Then, as the level of professionalism of the aesthetics increases from D1 to D5 the perception of bias in the news article returns to the D0 ground trust

bias ratings. Labels have been added to the top of the graph to show what comparisons will be undertaken to satisfy the H1 and H2 experiment hypotheses.

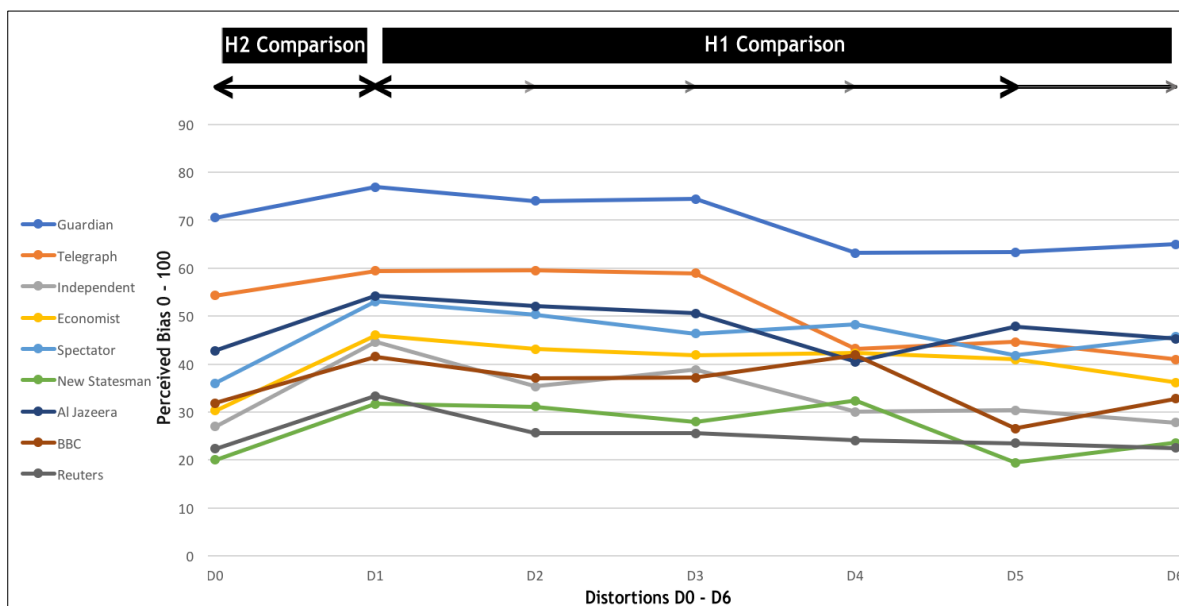


Figure 5-5 A graph showing the Mean bias scores for the nine distorted webpage/article combinations in Table 5-5.

Figure 5-6 depicts the webpage/article combinations D0 to D6 of the Telegraph, the Spectator, and the Independent from Figure 5-5 above. The graph lines for each are use the same colour schemes. Also marked are the results with significant differences. These are indicated with square boxes on each line. As highlighted earlier, one significant difference was found between D1 and D6 on the Telegraph. One significant difference was found between D0 and D1 in the Spectator. Two significant differences were found between D0 and D1 and D1 and D6 on the Independent. These results are also shown in Table 5-5.

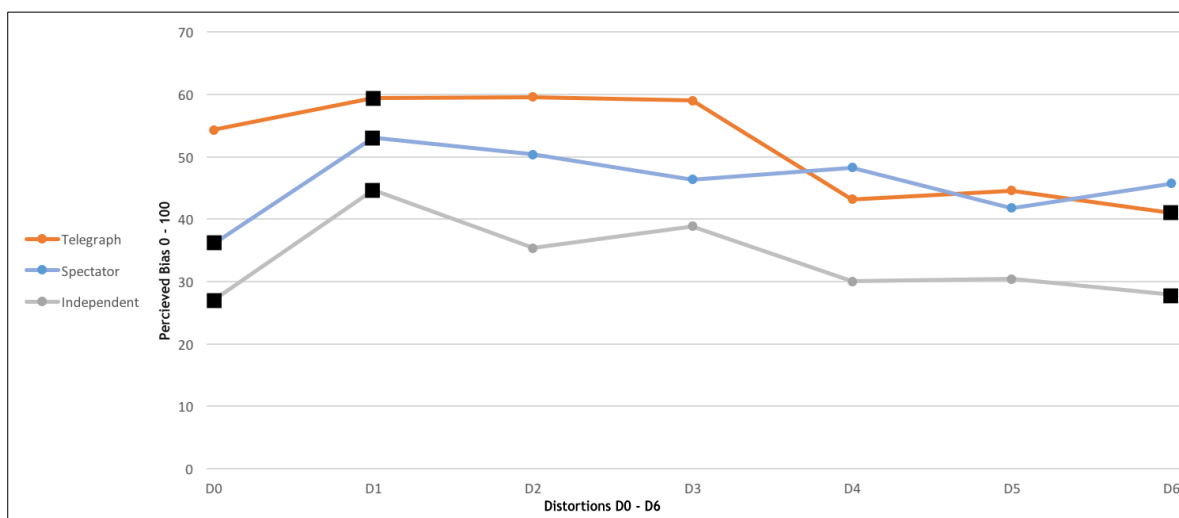


Figure 5-6 The Telegraph, Spectator and Independent webpage/article combinations from Figure 5-5. The square boxes indicate significant differences between distortions as per Table 5-5.

5.10.1 D0

The D0 column in Table 5-5 depicts the Mean bias rating for the plain text version of each webpage/article combination. This was measured to provide a ground truth bias rating for each article from the same population of the experiment participants. As was explained in section 5.7.1.2, the articles were modified to create a range from relatively unbiased to extremely biased. Each had a designed bias rating from 10.00 – 90.00, see Table 5-1. This was done to see if any effect was common across all levels of biased content. While this process was successful, the range is much more constrained than intended. The least biased article was rated at 20.02 and the most was rated as 70.51. It is possible that the pre-test participants knowing that the purpose of the process was to create a range provided a greater spread of results. However, as can be seen in Figure 5-5, a range of biased articles, though constrained, was achieved.

5.10.2 D1

The D1 column in Table 5-5 contains the Mean bias rating and Standard Error for the most degraded, least professional webpage/article combinations. As can be seen from Table 5-5 and Figure 5-5, these were considered the most biased. There was a marked increase from the D0 to D1, an average of 11.75% for the same articles, but displayed in the least professional aesthetic.

5.10.3 D2

D2 is a lightly less degraded version of D1, though it is still among the least professional of the webpage/distortion combinations. As can be seen from Table 5-5 and Figure 5-5, as the level of professionalism increases, the perception of bias from D1 to D2 decreases by an average of 3.61%. However, every webpage/article combination with the D2 level of degrading applied was considered more biased than their respective D0 ground truth.

5.10.4 D3

The D3 column in Table 5-5 contains the Mean bias rating and Standard Error for the slightly less degraded (compared to the D2) version of each webpage/article combinations. While there was some inconsistency, overall there was an average reduction in perceived bias of 0.73%. It is possible that the inconsistencies are due to the fact that the webpage/article combinations did not always degrade in a common manner due to the different use of underlying technologies. However, overall as level of degrading decreased and the level of professionalism in the aesthetics of each webpage increased, the perception of bias decreased.

5.10.5 D4

The D4 distortion is a less degraded version of the D3, and other than the D5, which is just an unbranded version of the undistorted D6, it is the least distorted webpage/article combination. The Mean bias ratings

and Standard Error for these are contained in the D4 column of Table 5-5 and the results are also depicted in Figure 5-5. While there was some inconsistency, as can be seen from the results, the closer each webpage/article combination came to its original and most professional form, the perception of bias was reduced by an average of 4.01%. This was especially noticeable in two of the three most biased articles displayed in the Guardian and Al Jazeera.

5.10.6 D5

The D5 distortion is simply an unbranded version of the D6. It is the aesthetically most professional version of each webpage/article combination with the name and logo replaced with the generic NewsCom moniker and a circle logo in the same colour scheme as the webpage. As the effect of the cumulative distortions was removed, and the level of professionalism returned to normal, the perception of bias in the articles dropped by an average of 3.02% further demonstrating that the more professional a news article is presented the less biased is perceived by its reader.

5.10.7 D6

The D6 distortion is an original version of each webpage/article combination with the branding included. It is not used to derive any findings and was only included to complete the offering and to show that the D5s bias rating scores were roughly on par with an average reduction in bias from D5 of just 0.16%. This demonstrates that the most professional D5 distortion without logo or branding is a good reflection of the original webpages.

Of note is the fact that the bias ratings for both D5 and D6 are roughly equal to those for their respective D0s, showing that the professional presentation of news articles only equals a plain text presentation in experiment conditions. It is unknown whether presenting a news article in plain text online would also be considered as professional as presenting it within a professional design.

5.11 Findings from Experiment Results

5.11.1 Comparison of D0 Ground Truth to D1 Least Professional

For each website in Table 5-5, D0 is the control, a plain text version of each article and headline with no design applied. These provide the ground truth bias rating for each article. D1 is the most degraded, least professional version of each webpage/article combination for each website. As can be seen from Table 5-5 and Figure 5-5, the perception of bias within each news article increased markedly for every webpage/article combination. The average increase was 11.75%. While this increase is both large and common, it also was significant in two incidences. The increases in perceived bias from D0 to D1 in both the Independent and the Spectator was significant. Combined with the large and common increase across

the board shown in Table 5-5, and depicted in Figure 5-5, it is possible to claim that the unprofessional display of news articles increases perceived bias when compared to plain text versions of the same articles.

5.11.2 The Cumulative Impact of Professional Webpage Aesthetics: D1 to D5 and Vice Versa

The impact of increasingly professional webpage aesthetics is evident in the results of the sample main effects of distortions statistical analysis. These results are shown in Table 5-5, and depicted in graph form in Figure 5-5 and Figure 5-6. While there were some slight individual inconsistencies, overall the perception of bias decreased from D1 to D5, 3.61%, 0.73% 4.01%, and 3.02%. Combined, the 11.37% decrease in perceived bias was virtually identical to the 11.75% increase from D0 to D1. Although the ANOVA showed no significant results between the D1s and D5s, the effect was common across all 9 websites, categories of websites, and articles with different levels of bias. Based on these results and findings, this experiment demonstrated that as the level of professionalism in the aesthetics of a news webpage increases, the perception of bias decreases, and vice versa.

5.11.3 The Impact of the Amount of Bias in the News Articles

The impact of increasing and decreasing professional aesthetics on news articles with different amounts of bias can be broken down in four ways.

From Figure 5-5 and Figure 5-6, it is clear that in the articles with the most bias, the impact of the least professional D1 distortion is muted. This is visible in the much smaller increases in perceived bias from the D0 to D1 distortions for the webpage/article combinations for the Guardian and the Telegraph which had the most biased articles. This would suggest that it is harder to make an already biased article appear more biased by its presentation.

Conversely, it also appears that it is easier to make news articles with less bias appear more biased due to unprofessional presentation. This is evident in Figure 5-5 and Figure 5-6, where the website with the webpage/article combinations with the most amount of bias registered the greatest increase from D0 to D1. This lends weight to the argument that mainstream or quality news organisations who claim to disseminate unbiased news have the most to lose from unprofessional presentation of news articles.

Most worryingly, the results of the experiment demonstrate that aesthetically professional webpages have the greatest impact on reducing the perception of bias in the most biased news articles. This is evident in the large reductions in perceived bias which the two most biased webpage/article combinations received as the level of professionalism increased. It should be noted that it is increasingly likely in the later less impactful distortions, D3, D4, and D5 that some participants recognised the webpage/article combinations as belonging to quality news agencies and therefore perceived them as less biased.

Lastly, the impact of professional presentation on reducing the perception of bias in the webpage/article combinations was consistent on articles with reduced amounts of bias.

5.11.4 The Impact of Category of News Website

Any impact of the category of news website, whether traditional print, news magazines, or international multi-format news agencies could not be determined from the results of the experiment. This was due to the use of news articles with different levels of bias which would make comparisons between websites difficult. The advantages and disadvantages behind the choice to include news articles with different levels of bias in the experiment are discussed in section 5.7.1.2.

5.12 Hypotheses Results

5.12.1 Primary Hypothesis

H₀ Perceived bias will not be decreased due to an increase in the professionalism of the news webpage's aesthetics.

H_A Perceived bias will be decreased due to an increase in the professionalism of the news webpage's aesthetics.

Based on the statistical analysis in section 5.9, the results presented in section 5.10, and the findings presented in section 5.11, the null hypothesis can be rejected and the alternative accepted. Perceived bias in news articles is decreased due to an increase in the professionalism of the news webpage's aesthetics.

This is evident by comparing the results of the D1 to the D5 distortions in Table 5-5, and the visible reductions in the perception of bias evident in Figure 5-5 and Figure 5-6. While there were no significant results between D1 and D5, there were two when comparing D1 to the D6 distortions, the bias scores of which were very similar to the D5 results. It is also worth noting that if the D1 distortions were made more unprofessional, there would likely have been several significant results proving the impact of professional aesthetics.

H₀ Perceived bias will not be increased due to a decrease in the professionalism of news webpage's aesthetics.

H_A Perceived bias will be increased due to a decrease in the professionalism of news webpage's aesthetics.

Based on the statistical analysis in section 5.9, the results presented in section 5.10, and the findings presented in section 5.11, the null hypothesis can be rejected and the alternative accepted. Perceived bias in news articles is increased due to a decrease in the professionalism of the news webpage's aesthetics.

Comparing the bias ratings of the plain text D0 distortions to their respective D1s clearly demonstrates a marked rise in perceived bias when the same plain text news article is presented in an unprofessional manner. This effect is common across all 9 news websites, sub-categories of websites and news articles with different levels of bias. On average, each webpage/article combination perceived bias increased by 11.75 points between its D0 and D1 distortions. The results are also backed up by two significant results, the Independent D0-D1 and the spectator D0-D1, see Table 5-5 and Figure 5-6. Consequently, we can reject the null hypothesis and accept the alternative hypothesis.

As previously stated in section 5.6.1.1, the two primary hypotheses are not simply opposites of each other, as it is possible that bias could only be affected in one direction.

5.12.2 Secondary Hypotheses

H₀ There will be no significant difference in the perception of bias between professional and unprofessional versions of the same news website.

H_A There will be a significant difference in the perception of bias between professional and unprofessional versions of the same news website.

The analysis presented in section 5.9, the results shown in section 5.10, and the findings derived in section 5.11 do not show any significant differences between the least professional D1 distortions and the most professional, though unbranded, D5 distortions. However, there are two significant differences between the least professional D1 and the most professional, branded D6 distorted webpage/article combinations for the Telegraph and the Independent, see Table 5-5 and Figure 5-6.

However, as there are no significant results between D1 and D5 and only two between D1 and D6, the findings are not sufficient to categorically reject the null hypothesis, and the alternative hypothesis therefore must also be rejected. It is worth noting however that if the webpage/article combinations had been distorted further, it is likely that significant differences would have been found.

H₀ The impact of the professionalism of the aesthetics will be inconsistent, regardless of the level of bias in the article.

H_A The impact of the professionalism of the aesthetics will be consistent regardless of the level of bias in the article.

Based on the statistical analysis in section 5.9, the results shown in section 5.10, and the findings derived in section 5.11 it is clear that the impact of the level of professionalism in the aesthetics of a news webpage is consistent across news articles with different levels of bias.

While the impact of unprofessional aesthetics being muted, and the impact of professional aesthetics being slightly exaggerated on the most biased webpage/article combinations, overall the effect of each level of

distortion was fairly consistent. This is most evident in the graphs presented in Figure 5-5 and Figure 5-6. Consequently, it is possible to reject the null hypothesis and accept the alternative, the impact of the professionalism of the aesthetics is consistent regardless the level of bias in the news articles.

5.13 Design Implications

With the current focus on the quality and veracity of news, especially online, news website designers, online editors, and journalists, need to be fully aware that any reduction in the level professionalism of the aesthetics of a news webpage, can increase the perception of bias in the news article. The issues introduced in this experiment to reduce the professionalism of a news webpage, such as issues with alignment, obviously broken underlying code, broken features, and a reduction in the visual quality and placement of images, though uncommon, can occasionally be seen in the webpages of mainstream or quality news agencies.

These can be introduced in five ways.

- Either through the original design, though this is unlikely in most professional news organisations.
- Second, through bad technical implementation. This often manifests itself through lack of robustness for different types and formats of content, browser support, different screen resolutions, or poorly implemented layouts.
- Though lack of oversight or checking of conformance to content posting guidelines in day-to-day use. Journalists desire to get their articles read and noticed by being on the front page of newspapers, the main story in television news shows, or onto the homepage of news websites has been long documented. Baron maintains: *“journalists may bias their stories if their career prospects can be advanced by being published on the front page”* (Baron, 2006). While Zaller wrote: *“In the case of journalists, career success means producing stories that make it onto the front page or get lots of airtime on the evening news, from whence flow fat salaries, peer respect, and sometimes a degree of celebrity status.”* (J. Zaller, 1999). Consequently, the journalists themselves are often responsible for pushing for bigger, more prominent, headlines, artwork, and eye-catching imagery. One of the visual distinctions between the most reputable quality press such as; Reuters, the New York Times, the BBC, and much less reputable sources such as; the Daily Mail, the Drudge Report, Mother Jones, the National Enquirer, is the rigidity and conformance to strict publishing guidelines. These control the number of elements on a homepage, their size, prominence or intensity, the strict use of headlines, colours and fonts, and the content of images, especially in relation their invasiveness, sordidness, and/or salaciousness. The less reputable, more biased publications tend to have a much broader mix-match of competing elements on their homepages or internal webpages which conveys unprofessionalism to the user.
- Through poorly integrated third-party content or services such as an abundance of low quality, often intrusive advertising, or social media. In some instances, these may display low quality

content which reduces the professionalism of the webpage's overall visual aesthetic. In other instances, they are obviously and visually broken or showing no content.

- The most reputable news agencies spend large amounts of money not only on high quality technical implementations of existing technologies, but also adopting and the appropriate use of emerging new technologies as solutions to display content. However, the principle of *primum non nocere*, first do no harm, must take precedence. While skilful implementation can act to differentiate them from news organisations with a smaller budget and/or less technical skills, both indicative of lower levels of professionalism, a bad implementation has the opposite effect. A similar effect has been seen on financial websites where issues with the design decreased credibility (Allport & Pendley, 2010; W. S. Jung et al., 2017).

While more investigation in the domain is required, the work of Alsudani and Casey highlighted 'Unity' in design between elements such as balance, harmony, contrast and dominance were instrumental to immediate judgements of credibility of which bias is a core dimension and measure (Alsudani & Casey, 2009). While the underlying technical and structural features are important, a seamless, professional design is required to tie all the elements together in order to create professional aesthetics on individual webpages. This is especially important in online news where the competition for audience share is extreme and there are no barriers to users moving between services.

Decreasing the professionalism of the website's design or a webpage's visual aesthetic, and thus increasing the perception of bias in the news articles they contain has long term serious consequences for news organisations. As Fico et al. has previously demonstrated, as perceived bias increases, perceived credibility of a news agency decreases (Fico et al., 2004). Thus, users are less likely to believe the articles they read are a fair and impartial telling of the facts or a balanced analysis of events, and they are therefore more likely to turn to one of the other easily available sources of news online.

5.14 Implications for News Consumers

The core concern for news consumers raised from the findings of this study is that biased news could be perceived as being less biased by presenting it in an aesthetically professional manner. Based on the statistical analysis in section 5.9.3, the results shown in Table 5-5, which are depicted in graph form in Figure 5-5 and Figure 5-6, the findings derived from them in section 5.11, and the results of the hypotheses in section 5.12, the impact of increasingly professional news webpage aesthetics on reducing the perception of bias in news articles with a range of biased content is clear.

As demonstrated in the introduction to bias in section 2.3.1 of the State of the Art chapter of this thesis, claims and counter claims of bias in the news are as old as the construct itself. While many believe that the news has become more biased in recent years, the evidence does not support this. News is biased. Or as

Weaver commenting on Efron's early study demonstrating this fact from the early 70s stated "*In such an age, the proposition that TV News is biased should be about as controversial as the law of gravity.*" (Efron, 1971; P. H. Weaver, 1972). Even the most high-quality news produced by the most reputable journalists and news organisations is susceptible to selection bias or gatekeeping bias at the very least. Or as D'Alessio and Allen quoting Debenport state: "*Journalists are aware that it is impossible to be absolutely objective and unbiased at all times (Debenport, 1992) and so attempt to substitute specific journalistic goals such as 'fairness' or 'balance.'*" (D. D'Alessio & Allen, 2000; Debenport, 1992).

Ironically, the well-known cognitive bias, called the Illusory Superiority bias, a common effect of which is that the majority of drivers believe that they are above average, which makes many think that the news they access is unbiased while the news that others access is biased (McCormick et al., 1986; Svenson, 1981). News consumers need to be aware that the information and messages in the news that they consume is being influenced by its presentation. While this study demonstrated the impact of professional aesthetics on the perception of bias in news websites, other dimensions and measures of credibility such as; accuracy, fairness, balance, expertise, and depth of coverage could similarly be influenced while consuming news or other forms of information in other domains.

5.15 Discussion

This research was undertaken under the supposition that a news website's design and individual news webpage's aesthetics impact the perception of bias within a news article. This was based on the knowledge that a website's design and its structural features have previously been shown to impact perceived credibility, of which bias is a core dimension and measure, especially when judging the credibility of news (Fogg et al., 2003; Spillane et al., 2018).

This study has proved conclusively that perceived bias in news websites, like credibility, is at the very least, partially judged heuristically. This finding has long term ramifications for the domain, especially considering how little research has been done on bias in the dissemination of news online. This disparity is evident in the comparatively small amount of relevant studies identified investigating bias in the dissemination of news online which are shown in Table 2-3. This research may help to begin to address this imbalance.

Another critical contribution absent from the body of knowledge is the lack of a human centric model explaining how judgements of bias are formed, online or otherwise. Such contributions are important as they form the theoretical underpinnings subsequent empirical work is based on. However, this situation can almost be described as a causality dilemma. Such theories are typically underpinned by an existing body of empirical work which they were derived from. Once developed, they in turn become the underpinnings of subsequent studies in a continual process of refinement as has happened with the ten,

human centric MTFS for judgements of credibility. While no significant body of work exists investigating how judgements of bias are formed online, it is unlikely that any strong theories will be developed to explain it. What is surprising about this situation is the lack of historical theory explaining judgements of bias in traditional mediums. The closest is Zaller's 'two message model', however this is too focused on decisions based on cues within the message in political communication (J. Zaller, 1999). Judgements of bias are not limited to politics, or as this study has demonstrated, even limited to the message. They may also be made without exposure to a message based on preconceived impressions of the source or medium. This is a definite lacuna in the research and should be addressed in the future.

Traditionally the consumption of news was a definitive act in space and time. Watching the 6 o'clock or 9 o'clock news, or reading the morning paper. Consumers were limited in the news they were exposed to. Whether it was by the papers they selected, or which their local shop stocked, or the news which the editors and journalists covered, or their access beyond local and a limited number of national television stations. News consumption was limited to local events or national events of importance which trickled down to available mediums and formats. In comparison, consumers are exposed to an almost unlimited range of local, national, and international news, in today's world. While many local newspapers failed, they have been replaced by additional radio stations and television channels, many of which now offer always on, 24-hour news, with headlines at 15-minute intervals. Not only do phones enable access to an almost unlimited range of news websites from around the world, but because of push notifications, the consumer is alerted within minutes of an important event or a news story relating to a topic which they have expressed interest in previously. As a result, many news consumers do not have the time to properly read and digest this constant stream of updates. News is scan read, or headlines and notifications are perused and dismissed with impunity. Many news consumers visit news websites for minutes or less at points throughout the day to get a sense of events. This rapid consumption provides the consumer with a visual snapshot of what is being discussed, what is important. Depending on how much time they have they may take a more in-depth scan of the headlines, and the imagery, while scan reading the teaser or lead text of each story on the homepage to get a sense of how the events are being treated. As a result, users are more exposed to the influences of the effect of visual biases in the presentation of news. This phenomenon is virtually unexplored.

Another unexplored concern is the fact that different types of design and aesthetics may instil a sense of professionalism, thus reducing the perception of bias and increasing the perception of credibility for different types of users. This is also an avenue worthy of future exploration.

Lastly, a large volume of studies, both from an academic and a commercial prospective, utilise A/B testing to the detriment of alternative forms of investigation such as qualitative or mixed methods approaches investigating e.g. the factors behind judgements, predisposition, heuristics, or context. One reason for this

is that there is an established bank of statistical tests for A/B testing and similar quantitative investigations. Some may feel that the ‘answer’ provided by such tests is easier to understand, to get published, or to convince a client to pay for, compared to qualitative or mixed methods investigations. This has resulted in a significant gap in the body of knowledge which is worthy of investigation.

5.16 Why is this Happening?

Since the earliest studies investigating the credibility of a source as a communicator of a message, researchers have found significant factors surrounding the message which increase or decrease perceived credibility. In television, these dissemination factors include physical characteristics such as gender, age, physical presence, and even their hair and facial expressions. This study, and the previous experiment, have demonstrated that users at least in part, use visual cues to make heuristics judgements via factors such as the level of professionalism in a website’s aesthetics and overall design, to make judgements of bias online.

5.17 Reflective Questions

As part of this experiment, a series of reflective questions were asked on a range of topics including; the experiment the participants completed, news website professionalism, bias, and credibility. A full breakdown of each of the responses is available in section 9.22 of the appendix. The following is an overview of the most important responses.

5.17.1 Recognition of Experiment Content and Perception of Bias

Table 5-6 depicts the results of two questions designed to ascertain whether the websites used in the experiment were recognised by the participants and whether they considered any to be particularly biased. Both the Guardian and the BBC were the jointly most recognised which was to be expected based on their popularity. However, the BBC was considered among the least biased, while the Guardian was considered among the most. However, it is possible that the results of this were influenced by the fact that participants had just completed an experiment where the Guardian was paired with the most biased news article and the BBC was paired with the fifth most biased. This is detailed in section 5.7.1.3, shown in Table 5-1, and depicted in Figure 5-1. However, as there were several reflective questions on different but related topics between this question and the experiment, and because the question specifically points to news websites, after they have been asked if they recognise any of these news websites outside of this experiment, the results should be treated as participant’s perception of how biased the news websites are in general.

The most interesting finding from this is that more people considered Al Jazeera biased than recognised it. One possible reason for this is the fact that it has an Arabic name.

Table 5-6 Results of two reflective questions asking participants which news websites they recognised in the study and which they thought were biased.

Question: Did you recognize any of the below News websites used in this experiment? Select each that you recognized.				
Question: Do you consider any of these News websites used in this experiment to be especially biased? Select each that you consider biased				
Website	Recognised	% 405	Biased	% 405
Guardian	207	51.11	118	29.14
Telegraph	122	30.12	75	18.52
Independent	88	21.73	60	14.81
Economist	152	37.53	61	15.06
Spectator	44	10.86	40	9.88
New Statesman	33	8.15	29	7.16
Al Jazeera	117	28.89	140	34.57
BBC	207	51.11	40	9.88
Reuters	143	35.31	29	7.16
None	49	12.10	77	19.01

5.17.2 Impact of Design and Aesthetics on the Perception of Bias and Professionalism

Four questions were asked about the impact of a news website’s design and webpage aesthetics on the perception of bias and professionalism, the results of which can be seen in the four tables in section 9.22.1. The results show that a news website’s design and aesthetics play a significant role in consumers' perception of bias and professionalism in both the news article and in the news organisation behind it.

- When asked whether design and aesthetics could impact the perception of bias in a news article, participants responded: Yes 31.60%, No 24.44%, Maybe 39.01%, and I Don’t Know 4.94%.
- When asked whether design and aesthetics could impact the perception of bias in a news organisation behind the website, participants responded: Yes 38.02%, No 25.68%, Maybe 31.36, and I Don’t Know 4.94%.
- When asked whether links to supporting content could impact the perception of bias in the organisation behind the website, participants responded: Yes 51.85%, No 20.99%, Maybe 23.95%, and I Don’t Know 3.21%.
- When asked whether links to supporting content could impact the perception of professionalism in the organisation behind the website, participants responded: Yes 65.92%, No 14.56%, Maybe 17.53%, and I Don’t Know 1.98%.

5.17.3 The Identification of Features and Characteristics of a News Website's Design, Reflected in a News Webpage, that Impact Perceived Professionalism and Bias

In two questions, participants were also asked which features and characteristics of a news website's design, reflected in a news webpage's aesthetic, or the articles contained therein, impact their perception of bias and professionalism. The results are in two tables in section 9.22.2 of the appendix. For both questions, participants were presented with the same 28 multiple choice options selected based on the findings of the previous experiment, from the literature, and a review of popular news websites. 14 of the options focused on the text of the article and the message contained therein, while the other 14 focused on its presentation.

5.17.3.1 Professionalism

The results of this validate the process described earlier to degrade the professionalism of the webpage/article combinations. While 'Bad quality writing' was selected by 76.05% of responders as impacting perceived professionalism, 70.86% selected 'Low quality advertisements', 64.94% selected 'Too many advertisements', and 55.06% selected 'Advertisements which are too prominent or gaudy'. In fact, of the total participants' responses for all 28 text and presentation features and characteristics, 2953 were for text, while 2417 were for presentation. This demonstrate the significant impact that a website's design and a webpages visual aesthetics have on the professionalism, see Table 9-38.

5.17.3.2 Bias

Unsurprisingly, when participants were asked which features and characteristics of a website's design, reflected in a news webpage's aesthetic, or the news article therein, impact their perception of bias, the majority focused on issues with the text such as 'Unsupported claims in the article' 60.25%, Lack of objectivity in the article' 56.30%, or 'Omitted facts from the article' 54.07%. However, two elements focused on the presentation did make it into the top ten. 'Bad or broken alignment in the design' 50.12% and 'Low quality images' 49.14%. In total, the 405 participants selected text focused features and characteristics 2630 times, while they selected presentation focused features and characteristics 2206 times, see Table 9-38.

5.17.4 Most Credible and Most Biased News Mediums

Participants were also asked which the most credible and most biased news medium was. The results of these are provided in Table 9-40 and Table 9-41 in section 9.22.3 of the appendix.

5.17.5 Most Relied on Dimensions of Credibility When Judging News Online

The final question asked of participants was to select, via multiple choice, which dimensions of credibility they rely on when judging news on the WWW, see Table 9-42 in section 9.22 of the appendix. The 25

dimensions were selected based on the most common measures identified in the online repository and classification of credibility measures³⁸ compiled as part of the CAFE framework which was detailed in section 3.6.3. This question was asked to back up previous work by Fogg et al. who demonstrated that 11.6% of participants in a large study cited information bias as a particular concern when evaluating the credibility of information online (Fogg et al., 2003). However, this rose to 30.2% when evaluating the credibility of news online. While bias was considered an important measure by 53.58% of participants, an even higher percentage than the study by Fogg et al, six other measures were considered even more important. Accuracy 67.9%, Professional 59.51%, Honesty 57.78%, Depth of Coverage 57.53%, Factualness 55.06%, and Objective 54.07%.

5.18 Lessons Learned from Experiment Two

There was one important lesson from Experiment Two:

- Based on the findings of the news website design review process and the responses of participants in the reflective questions in both experiments, there is little doubt of the impact of colour schemes on the perceived level of professionalism in news websites. However, no distortion was applied to them to degrade their professionalism due to issues with repeatability and subjectivity. In hindsight, this was a mistake.

Issues with repeatability could have been solved by recording the Red Green Blue (RGB) or the Cyan Magenta Yellow Key (CMYK) of the original colours and the alternative colours as they were substituted during the distortion process. Issues with subjectivity could also have been solved by pre-testing the original professional colours applied to each webpage and a range of alternative unprofessional colour schemes applied to each webpage design. Using this process, it would have been possible to identify a suitable alternative unprofessional colour scheme for each webpage included in the experiment.

This was an oversight as it could be argued that without degrading the colour scheme for each webpage/article combination, the least professional distorted webpage/article combination did not acutely or fully reflect the aesthetics of an unprofessional news webpage. Consequently, it is possible that the perceived bias scores for the D1-D4 distorted webpage/article combinations could have been worse and likely would have resulted in additional significant results. This should be accounted for in future.

³⁸ https://www.scss.tcd.ie/~spillab/Credibility_Measures/

5.19 Summary of Experiment Two Findings

There are two main findings from this experiment. First, as the level of professionalism in a news webpage aesthetic decreases, the level of perceived bias in the news article increases. Second, and conversely, the opposite is also true, when the level of professionalism in a news webpage increased, the level of professionalism decreases. Third, though there were some slight differences, see section 5.11.3, these effects were broadly constant across webpages displaying articles with different amounts of bias.

Consequently, it is possible to claim that perceived bias in a news article online, is at least partially, judged heuristically based on the webpage's design and aesthetics. These findings are consistent with previous studies on credibility, of which bias is a core dimension and measure, especially when judging news online.

The reflective questions stage of the experiment also confirmed and revealed a number of important findings. These include the importance of website design and webpage aesthetic factors to judgments of bias, professionalism, and credibility online, both in the news article and in the organisation behind the news website. It also revealed the most and least credible news mediums and demonstrated that increasing numbers of younger users rely on social media to access news because it is perceived as being more credible. Confirming the importance of bias as a dimension of credibility, especially when judging news online, the participants in this experiment also considered social media to be the least biased medium by an overwhelming margin.

5.20 Conclusion

This study satisfied two research objectives outlined in section 5.3.1, and partially satisfied the major research contribution, outlined in section 5.3.2. The experiment was designed with an early version of the CAFE framework and the experience gained from this experiment was used to refine its development. It was also deployed on the accompanying experiment platform, the codebase of which was significantly updated and upgraded. It demonstrated that as the level of professionalism in the aesthetics of a news webpage increases, the level of perceived bias in the news article decreases. The opposite is also true, as the level of professionalism decreases the perception of bias increases. This is most obvious in the comparison of the D0 ground truth bias ratings for each webpage/article combination to their respective D1 least professional webpage/article combinations.

The major contribution of this experiment to this thesis is the identification of professionalism as a characteristic of the design and aesthetics of news webpages which can influence perceived bias in the news article contained therein. This is an important contribution as it means that there could be unknown or undue influences on the perception of the information in news articles of the message contained therein. It also partially answers the first part of the research question by demonstrating that characteristic of a news website's design, reflected in a news webpage's aesthetics, can influence users' perception of bias.

Furthermore, it lays the foundation for investigating and answering the second part of the research question, namely, whether the deliberate adaptation of such features or characteristics can be used to predicatively influence the perception of bias. Due to the strength of the findings from this experiment, it was decided to focus on characteristics.

“The best argument against democracy is a five-minute conversation with the average voter.” Winston Churchill

6 Experiment Three – Increasing and Decreasing the Perception of Bias Through News Website Design and Webpage Aesthetics

6.1 Introduction

As has been established as part of this research and previously, bias is a core dimension and measure of credibility, especially when users are judging news online (Fogg, 2003; Spillane et al., 2018). The previous studies have also demonstrated that like credibility, perceived bias in the content of a message, in this case a news article, is affected by its visual presentation, proving that heuristics play a part in judgements of bias. The objective of this experiment is to evaluate whether the deliberate adaption of a characteristic of a website’s design, reflected in a webpage’s aesthetic can be predictively impact the perception of bias. the resulting contribution forms the second half of the research question. These are further detailed in section 6.3. Consequently, it was decided to focus on the characteristic of ‘High Quality Aesthetics’ as participants in the first experiment reported in the reflective questions that it was the most influential visual characteristic to reduce perceived bias, see section 4.14.3.3 and Table 9-28 in the appendix. The quality of the visual aesthetics also separates many of the different categories of news website.

6.2 Influence from the State of the Art and Experiments One and Two

The reflective questions from experiment one revealed four characteristics of a website’s design or a webpage’s aesthetics which participants felt reduced the perception of bias within a news article. The first two, Seriousness and Professionalism, are considered primarily characteristics of source or the organisation behind the website. Professionalism rather than seriousness was investigated in the second experiment for two reasons. First, because they are closely related, and second, because there is an existing body of literature on professionalism. This experiment, therefore, focuses on the closely related presentation characteristics of high-quality aesthetic and good design.

High-quality aesthetics and good design are distinct from the related characteristic of professionalism. In experiment two, professionalism was investigated by slowly and cumulatively lowering the technical and visual professionalism of the implementation to reflect an amateurish, or broken presentation. It was therefore possible to compare ground truth bias rating for each article with the least professional distorted webpage/article combinations, and to compare these to unbranded versions of the original professional implementations.

The quality of the aesthetics and the design quality are about the visual experience rather than the level of professionalism of the implementation. It is possible to have a professionally implemented low-quality news website or webpage. Therefore, for this experiment the distortions will concentrate on the increasing or decreasing the amount, type, and prominence of advertisements, and the quality of the product or service they display, reducing or increasing the prominence of calls to actions and their placement in the webpage, and the proportion of each page that is dedicated to the display of informative news content. Though out of scope for this thesis, there is a significant gap in the literature as to the commercial impact of advertisements on the perception of professionalism in news websites. One area of particular interest is predicting the impact of various types of advertising content, taking into account factors such as size, prominence, position, animation etc, on the perceived professionalism of the news website. Adding this information to individual user models which news websites use for personalisation, would make it possible to identify the maximum tolerable limit for advertisements for a particular individual based on a maximum allowable negative impact on perceived professionalism. This would allow news websites to maximise commercial revenue without detrimentally impacting their brand.

Another reason high-quality aesthetics was chosen is because it is visually one of the most obvious differences between different categories of news websites such as quality press, tabloid, and the newer online only generation of news websites which range from quality or semi-quality news agencies to virtual content farms. These are discussed further in section 9.23 of the appendix. Thus, many aspects of the visual experience including the quality of their aesthetics are greatly improved.

Therefore, it was decided to investigate whether by adapting the quality of the aesthetics of existing news websites to reflect the aesthetics of quality news agencies and low-quality content farms news websites, whether it is possible to predictively influence the perception of bias in the news articles they contain. This was undertaken under the supposition that high-quality aesthetics would decrease the perception of bias while low quality aesthetics would increase it.

6.3 Experiment Objectives and Contributions

6.3.1 Objectives

This experiment fulfils the final research objective. It was conditional on the results of experiment two.

- **Objective 4:** To evaluate the adaptation of such to predictively impact the perception of bias.

6.3.2 *Contribution*

This experiment forms the second half of the Major Contribution, namely the validation through experimentation.

- **Major Contribution:** The identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

6.4 *Research Ethics Approval*

Research ethics approval for this experiment was granted by the Research Ethics Committee of the School of Computer Science and Statistics, Trinity College Dublin.

6.5 *Use of Website Titles to Describe Results and Findings*

As stated in section 4.5 and section 5.5, it should be noted that in this and the two other experiments described in this thesis, news website names and titles are used to describe results. Any results and findings are not reflective of the publication in reality, and should be attributed to the news articles, which were deliberately edited to create a range of biased content and were inserted into each webpage, and the distortions applied to them. It is no reflection of the news agencies. It may also be partially ascribed to participants recognising the news agency and their positive or negative feelings towards it, though every attempt was made to prevent this.

6.6 *Experiment Design*

This experiment was set up as a 9x4 within-subject incomplete counterbalanced measures design. It was designed and deployed using the CAFE experiment framework and opensource platform that was used in the previous experiments. Incomplete counterbalancing was achieved by arranging the webpages and distortions, D0–D3, in a reduced form Latin square. Nine news webpages, from popular news websites³⁹, were selected for inclusion in the experiment. This is described in detail in section 4.7.1.1. The original article in each webpage was removed and replaced with one of nine articles with different levels of bias, see Table 5-1 and Figure 5-1. Participants were randomly assigned to one of nine diagonal paths through the Latin square that intersected with each distortion, webpage or article, thus ensuring that each participant experienced each distorted and webpage/article combination once. To reduce carryover effects and task fatigue, once assigned to a path the webpage/article combinations the participant would encounter were displayed in random order. Two attention questions, to determine diligence were also added to each

³⁹ www.theguardian.com, www.telegraph.co.uk, www.independent.co.uk, www.economist.com, www.spectator.co.uk, www.newstatesman.com, www.aljazeera.com, www.bbc.com, and www.reuters.com

participant's path. 47 participants failed one or both of these attention questions and were not included in the final dataset.

This experiment was designed to confirm the findings of earlier experiments and to demonstrate that perceived bias could be predictively influenced by adapting characteristics, in this instance the quality of the aesthetics, of a website's design or a webpage's presentation. For efficiency, this experiment was conducted at the same time as experiment two. Therefore, certain aspects such as the selection of websites, webpages, content and participant numbers are the same.

6.6.1 Hypothesis

The central tenet of this research is that users' perception of bias within a news article can be predictively increased or decreased by distorting the quality of its visual presentation within a news webpage. To this end, three hypotheses were formed.

The first is based on the supposition that by reducing the quality of the visual experience, perceived bias within the article will be increased when each is compared to its respective control.

H₀ Perceived bias will not be increased due to a reduction in the quality of visual presentation.

H_A Perceived bias will be increased due to a reduction in the quality of visual presentation.

The second hypothesis is based on the supposition that by improving the visual experience that the perception of bias in each article will be reduced when each is compared to its respective controls.

H₀ Perceived bias will not be decreased due to an improvement in the quality of visual presentation.

H_A Perceived bias will be decreased due to an improvement in the quality of visual presentation.

The quality of visual presentation has previously been shown to impact credibility and some of its individual measures such as trust. This hypothesis aims to ascertain if the same is true for bias.

H₀ There will be no difference in the perception of bias between high-quality and low-quality versions of the same news article.

H_A There will be a difference in the perception of bias between high-quality and low-quality versions of the same news article.

6.7 Methodology

The experiment required participants to rate perceived bias in nine webpage/article combinations which were each subject to four distortions, thus resulting in a 9x4 within-subject incomplete counterbalanced

measures design. A single webpage from the nine aforementioned news websites was paired with a news article creating nine webpage/article combinations. These were then subject to four distortions. D0 is a plain-text version of the news article, D1 is an original version of the webpage/article combination with branding removed, D2 is a version with improved webpage aesthetics, and D3 is a version with degraded webpage aesthetics. For efficiency, this experiment was carried out at the same time as experiment two (Spillane, Lawless, et al., 2020), and thus there is some overlap in the webpage/article combinations they experienced, see Figure 6-1.

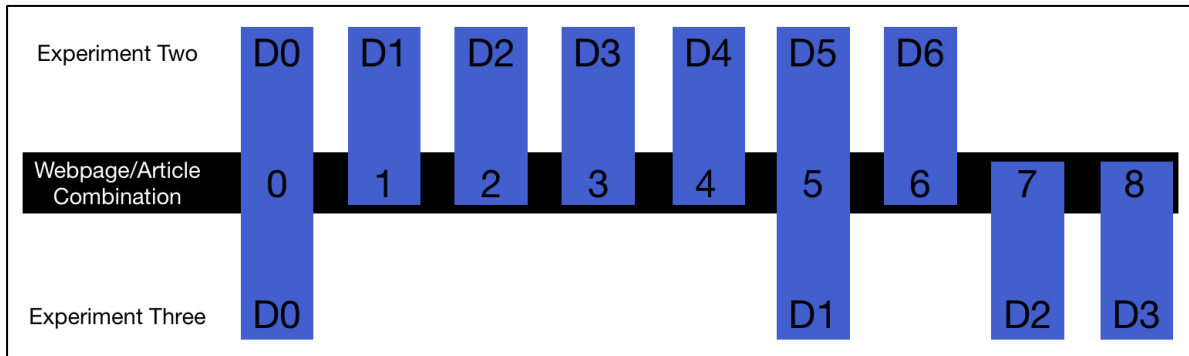


Figure 6-1 Breakdown of experiments two and three showing how they utilised the same webpage/article combinations but that for the most part different distortions were compared in each experiment.

6.7.1 Experiment Content

6.7.1.1 Websites and Webpages

A description of how the websites and webpages were selected is available in section 5.7.1.1 of the last chapter. In brief, a single webpage from nine popular news websites were selected for inclusion in the experiment. This was then paired with a news article for the duration of the experiment

6.7.1.2 Creating Biased News Articles

Section 5.7.1.2 of the last chapter describes how the news articles were selected and how they were modified to create a range of content from relatively unbiased to extremely biased. This was to discover if any effects were common across different types of news content. This process was subject to a pre-test to ensure the effectiveness of the procedure, the results of which are available in Table 5-1.

6.7.1.3 Pairing of News Webpages and News Articles

This process is described in section 5.7.1.3 of the previous chapter. The webpages each article was assigned to can also be seen in Table 5-1.

6.7.2 *High Quality Aesthetics and Good Design*

To understand which elements should be selected for distortion, and the form these distortions should take, a design review was undertaken, following the format of Zhang and von Dran, of the nine websites used in this experiment and those of other news organisations (Zhang & von Dran, 2000). Twenty-six features and characteristics of news website's design, reflected in a news webpage, which can convey quality to the viewer were identified. Using a factor analysis, these were then assigned to eleven categories. Features were defined as elements of a news website used to provide information or a service, e.g. search facility, latest news, or most read feature. Characteristics were defined as the abstract influences on the website's design such as the business, gaudiness or loudness of the design. The underlying means of affordance of each of these eleven categories was then identified. Sundar maintains that the credibility of a website is conveyed by visual cues in the design which are indicative of the underlying affordances in the technology behind the website (Sundar, 2008). There are two ways they convey cues. First, by their sheer presence, and second, by assembling information, related to the message or the topic, which he previously coined 'information scent', that is pertinent to the user (Sundar et al., 2007). Information scent was also used to categorize one of the groupings of identified features and characteristics in the factor analysis.

The D2 and D3 rows in Table 6-1 depict the eleven categories identified in the factor analysis. Nine of these were distorted to reduce or improve the visual quality of the webpages. Two categories, Name/Branding and Colour were not used to create the high quality (D2) or the low quality (D3) distortions. Name/Branding was not used as all names, branding and logos were replaced with the generic moniker, NewsCom with a plain circle logo in each webpage's colours in the D1 distortion, from which the D2 and D3 descend. The name, branding, and logos were not further distorted to create the high quality (D2) or the low quality (D3) distortions. Changes to each webpages colour scheme to increase and decrease quality were not implemented due to being too subjective. It would also likely result in too great a change to the design to make any findings useful. Figure 6-2 depicts the D0, D1, D2 and D3 versions of the same BBC webpage used in this experiment. In each case, it displays the same news article. D0 is the plain text version to establish ground truth bias ratings. The D1, D2, and D3 demonstrate how the fundamental webpage was distorted to create the high quality and low-quality experience.

Figure 6-2 depicts the D0, D1, D2 and D3 versions of the same BBC webpage used in this experiment. In each case, it displays the same news article. D0 is the plain text version to establish ground truth bias ratings. The D1, D2, and D3 demonstrate how the fundamental webpage was distorted to create the high quality and low-quality experience.

Figure 6-2 has also been divided into four parts for greater clarity. See Figure 9-10, Figure 9-11, Figure 9-12, and Figure 9-13 in section 9.23 of the appendix.

Table 6-1 The results of the news website design review and the resulting distortion process for each of the distorted webpage/article combinations.

Distortion	Design Review	Distortion Process
D0: Ground Truth	NA	Plain text, <h1> and <p> tags only with no CSS or any other form of styling or imagery
D1: Control	NA	The control. This is an original copy of the webpage with branding and logo removed and replaced with the generic moniker NewsCom. The assigned news article was inserted into the underlying code of each webpage.
D2: High Quality	<p>Article Strength: Add inline text links, article metadata such as time stamps, article correction and graph to the article</p> <p>Lineage / Information Scent: Add supporting Content links, part of an article series, part of a global news network</p> <p>Layout: Fix or improve any alignment issues and reduce gaps</p> <p>Advertising: Increase quality/status of product, and quality of graphics. Reduce prominence, gaudiness and size. Remove inline ads from content or beneath top navigation</p> <p>Social media: Reduce number and size of icons, improve graphics</p> <p>Author Experience: Add author name, detailed profile with picture, experience</p> <p>Sponsored Content: Remove or Reduce</p> <p>Frivolous Content: Remove or Reduce such as click bait links</p> <p>Numbers: Add article time-date stamp and market data widget</p> <p>Name/Branding: None</p> <p>Colour: None</p>	<p>Inclusion of limited and restrained advertising of high-quality products</p> <p>Increase in the proportion of the screen space dedicated to displaying the news article</p> <p>Addition of a detailed author profile complete with picture and bio with related experience</p> <p>A graph or other form of data visualisation relating to the content on each page</p> <p>The addition of elements which emit information scent, such as links to related articles, references or sources, see Table 6-1, row D2.</p>
D3: Low Quality	<p>Article Strength: Remove inline text links, correction and/or graphs</p> <p>Lineage / Information Scent: Remove supporting Content links, part of an article series, part of a global news network</p> <p>Layout: Break alignment and add gaps</p> <p>Advertising: Reduce product status, quality of graphics. Increase prominence, gaudiness and size. Place advertisements where they interrupt flow</p> <p>Author: Remove author name, profile information and picture, experience</p> <p>Sponsored Content: Increase amount / prominence. Reduce quality</p> <p>Social media: Increase number and size of icons. Increase prominence, reduce quality of graphics</p> <p>Frivolous Content: Introduce or increase click bate links</p> <p>Numbers: Add article time and date stamp and market data widget</p> <p>Name/Branding: None</p> <p>Colour: None</p>	<p>Increasing the size, amount of, prominence, and gaudiness of advertising</p> <p>Reducing the quality of the products or services offered in the advertisements.</p> <p>Advertisements were placed in positions for maximum visibility displacing news content, navigation or other useful services, interrupting the visual experience.</p> <p>The percentage of the screen dedicated to sponsored content was significantly increased.</p> <p>Advertisements were also inserted into the middle of each article to interrupt the reading experience.</p> <p>All links to supporting content, similar articles, or any indication of a history or experience of covering the same topic, so called information scent (Sundar et al., 2007), were removed.</p> <p>Author name, credentials, photo and any other profile information were removed.</p> <p>Each news article was also anonymized.</p> <p>All meta data, including time stamps, article corrections, syndication information was removed.</p> <p>Social media interaction facilities, previously retained in the D1, were moved to the top of the news article, made larger and a demanding call to action on the user added.</p> <p>Overall the percentage of the screen focused on delivery of the news was reduced.</p> <p>Each webpage went through the same distortion process. E.g. the same low-quality advertisements were added in the same position to each distorted webpage/article combination.</p>

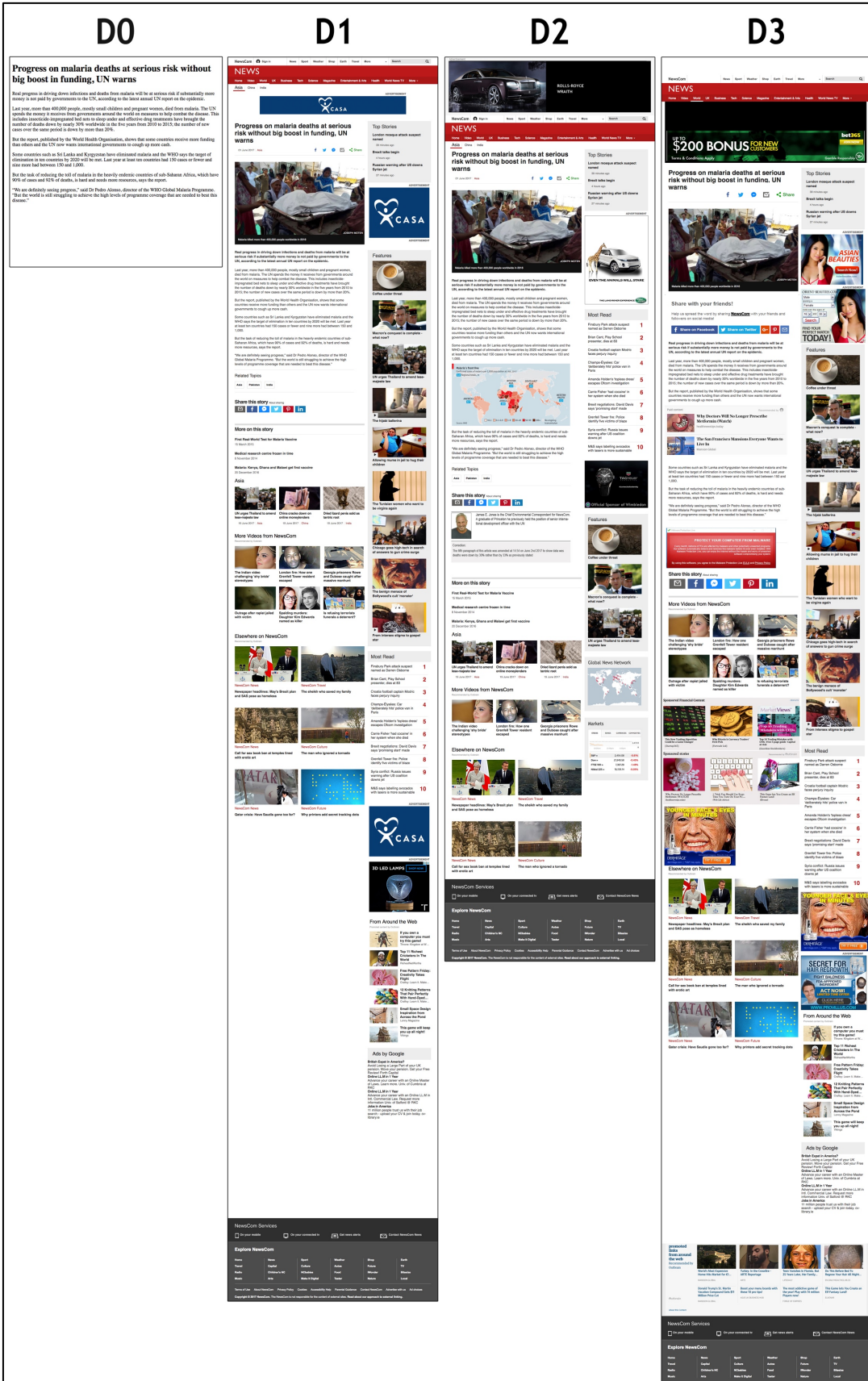


Figure 6-2 Four distortions of the BBC webpage/article combination. Note: see Figure 9-10, Figure 9-11, Figure 9-12, and Figure 9-13 in section 9.23 of the appendix for larger versions.

6.7.3 *Definitions of Bias and Participant Instructions*

Participants were then given the following definition of bias: “*Bias is defined as ‘Deliberate or accidental slant by the journalist, editor or publication to distort reality.’*” Participants were again presented with the above definition in a popup dialog during the first Instruction Task, and again after they answered the user profile questions. Section 5.7.6 of the previous chapter describes this in more detail.

6.7.4 *Crowdsourcing*

The process to describe the crowdsourcing of participants is described in detail in section 5.7.8 of the previous chapter. In brief, 405 participant submissions were included in the final dataset. All participants were over 18 and from the US with English as a first language.

6.7.5 *Validity, Statistical Power Analysis, and the Bias Rating Scale*

Section 5.7.9 describes in detail the six actions to ensure validity in the experiment setup. Statistical Power Analysis tests, described in detail in section 5.7.10, revealed a post hoc $>.95$ actual power to detect an Effect Size of 0.24. Finally, section 5.7.11 describes the benefits of using a VAS as a measurement device.

6.7.6 *Profile*

In total 405 submissions were evaluated. 51.4% Male, 47.2% Female, 1.4% Other. Mean age 36.48. 34% were <30 . Participants were spread across the US with 44 states and the District of Columbia represented. Education and Occupation revealed a highly educated, mostly professional sample. Further details of the profile of the sample population can be found in section 5.8 and particularly in Table 9-43 in section 9.24 of the appendix.

6.8 *Statistical Analysis and Results*

To derive the impact of the high quality and low-quality distortions, nine One-Way Repeated Measures ANOVAs using the Bonferroni correction were conducted as Simple Main Effects in SPSS 24. While Simple Main Effects would usually not be explored without a significant two-way interaction in a Two-Way Repeated Measures ANOVA, the focus of this research is on comparing the D2 and D3 distortions webpage/article combinations to their respective D1 controls and to each other. In SPSS, Simple Main Effects are undertaken by performing multiple One-Way Repeated Measures ANOVAs. Alternatively, one could perform multiple paired samples t-tests, however the increased validity of the One-Way Repeated Measures ANOVA due to use of Bonferroni correction is preferred. The data assumptions for Two-Way Repeated Measures ANOVAs are valid for One-Way Repeated Measures ANOVAs. Results of the nine, One-Way Repeated Measures ANOVAs using the Bonferroni correction are shown in Table 6-2. As the focus is on investigating the impact of distortions, Simple Main Effects for Webpage/Article combinations were not undertaken.

6.8.1 *Main Effects*

The experiment data had already passed the first two assumptions for ANOVAs, namely that the dependent variable was a continuous interval type data and the independent variable should consist of two or more independent groups. The data collected during the experiment was analysed for the remaining three data assumptions for ANOVAs. A visual inspection of boxplots showed a small number of outliers. It was decided to continue the analysis with the original data as the results would not be materially affected as there was little difference in the findings of the original and transformed ANOVAs. A Shapiro-Wilk test ($p = >.05$) for normality on the studentized residuals revealed that some of the data was not normally distributed. It was again decided to continue with the test as the One-Way Repeated Measures ANOVA is considered robust to deviations from normality when populations are >50 . This experiment had 405 participants. Due to the central limit theorem, One-Way Repeated Measures ANOVAs can still provide valid results even when the distribution of the data is very non-normal, when the population sample is ≥ 30 .

The following sections and Table 6-2 depict the results of the nine individual ANOVAs for each website. Although the D1 distorted webpage/article combination for each website is the control in this experiment, the D0 plain-text version was included in the analysis to provide ground truth bias ratings and a greater overview. Mauchly's Test of Sphericity (MTS) ($p >.05$) indicated that the assumption of sphericity had been met for each of the ANOVAs. Therefore, the Sphericity assumed significance values are reported without the Greenhouse-Geisser (GG) or the Huynh-Feldt (HF) correction.

The Two-Way Repeated Measures ANOVA revealed no statistically significant two-way interaction between webpages and distortions $F(24.00, 1056.00) = 1.049$, Sphericity Assumed $p = .399$. This was to be expected and the experiment was not designed to establish such.

The main effects of webpage/article combinations showed a statistically significant difference $F(7.11, 352.00) = 47.48$, Sphericity Assumed $p = <.0005$. Like experiment two, this was most probably due to the differing levels of biased content inserted into each webpage. However, as the focus is on investigating the impact of distortions, further analysis in the form of Simple Main Effects for Webpage/Article Combinations were not undertaken.

An analysis of the main effects of distortions showed that there was a statistically significant difference $F(3.00, 132.00) = 15.12$, Sphericity Assumed $p = <.0005$. These are discussed further in the following sections.

6.8.2 Simple Main Effects for Distortions

While Simple Main Effects for Distortions would usually not be explored without a significant two-way interaction, the focus of this research is on comparing the distorted webpage/article combinations to their controls and to each other. In SPSS, Simple Main Effects are undertaken by performing multiple One-Way Repeated Measures ANOVAs. Alternatively, one could perform multiple paired samples t-tests, however the increased validity of the One-Way Repeated Measures ANOVA due to the option of the Bonferroni correction is preferred. The assumptions required for Two-Way Repeated Measures ANOVAs are valid for One-Way Repeated Measures ANOVAs. It should also be noted that due to the central limit theorem, One-Way Repeated Measures ANOVAs can still provide valid results even when the distribution of the data is very non-normal. Table 6-2 shows the results of One-Way Repeated Measures ANOVAs – Simple Main Effects for Distortions, using the Bonferroni correction for multiple comparisons for each of the websites.

The Simple Main Effects for Distortions revealed five significant results. The colours correspond to the cells in Table 6-2 and the lines for each website in Figure 6-3.

• Guardian	D2 – D3	Blue
• Telegraph	D2 – D3 and D0 – D2*	Orange
• Reuters	D2 – D3	Dark Grey
• BBC	D2 – D3	Rust

*As the D0s are only included for ground truth bias ratings and are not part of the analysis for the hypotheses, significant difference between D0 and D2 for the Telegraph does not contribute to the experiment findings or answering the experiment hypothesis.

6.8.2.1 Guardian

The difference between D2 and D3 Mean bias ratings was statistically significant $F(3, 132) = 5.442, p = .002$, partial $\eta^2 = .110$. An analysis of the pairwise comparisons showed an increase in perceived bias from D2, 53.884 ± 31.577 to D3, 75.156 ± 18.619 (95% CI 5.207 to 37.415), $p = .004$. The D2 version of the Guardian, designed to have a high-quality visual experience, is considered less biased than its D3.

6.8.2.2 Telegraph

The difference between D2 and D3 Mean bias ratings was statistically significant $F(3, 132) = 5.856, p = .001$, partial $\eta^2 = .117$. An analysis of the pairwise comparisons showed an increase in perceived bias from D2, 35.773 ± 28.600 to D3, 58.420 ± 31.731 (95% CI 4.719 to 40.658), $p = .007$. A further inspection also showed a decrease in perceived bias from D0, 54.267 ± 27.908 to D2, 35.773 ± 28.600 (95% CI 2.005 to 35.062), $p = .020$. It should be noted that while D0 was included in the analysis, it is not used to satisfy any

of the hypothesis. The visually improved D2 version of the Telegraph was also considered significantly less biased than its diminished D3.

6.8.2.3 *Independent*

The difference between D2 and D3 Mean bias ratings was not statistically significant $F(3, 132) = 1.325$, $p = .269$, partial $\eta^2 = .029$. As no significant result was identified, a detailed analysis of the pairwise comparisons are not presented though F statistic and p values can be seen in Table 6-2.

6.8.2.4 *Economist*

The difference between D2 and D3 Mean bias ratings was not statistically significant $F(3, 132) = 1.965$, $p = .122$, partial $\eta^2 = .043$. As no significant result was identified, a detailed analysis of the pairwise comparisons are not presented though F statistic and p values can be seen in Table 6-2.

6.8.2.5 *Spectator*

The difference between D2 and D3 Mean bias ratings was not statistically significant $F(3, 132) = 2.671$, $p = .050$, partial $\eta^2 = .091$.

Too err on the side of caution, the p value of .050 was not counted as significant. However, its effect can also be seen in Table 6-2 where two of the Spectator's pairwise comparisons can be seen to be close to significant. As no significant result was identified, an analysis of the pairwise comparisons were not undertaken.

6.8.2.6 *New Statesman*

The difference between D2 and D3 Mean bias ratings was not statistically significant $F(3, 132) = .699$, $p = .554$, partial $\eta^2 = .016$. As no significant result was identified, a detailed analysis of the pairwise comparisons are not presented though F statistic and p values can be seen in Table 6-2.

6.8.2.7 *Al Jazeera*

The difference between D2 and D3 Mean bias ratings was not statistically significant $F(3, 132) = .858$, $p = .792$, partial $\eta^2 = .019$. As no significant result was identified, a detailed analysis of the pairwise comparisons are not presented though F statistic and p values can be seen in Table 6-2.

6.8.2.8 *BBC*

The difference between D2 and D3 Mean bias ratings was statistically significant $F(3, 132) = 3.506$, $p = .017$, partial $\eta^2 = .074$. Analysis of the pairwise comparisons showed an increase in perceived bias from D2, 23.24 ± 22.091 to D3, 39.44 ± 23.261 (95% CI 1.797 to 30.603), $p = .020$. The D2 version of the BBC,

distorted to improve its appearance, was rated as being significantly less biased than its visually diminished D3.

6.8.2.9 Reuters

The difference between D2 and D3 Mean bias ratings was statistically significant $F(3, 132) = 2.690$, $p = .049$, partial $\eta^2 = .058$. An analysis of the pairwise comparisons showed an increase in perceived bias from D2, 18.42 ± 19.262 to D3, 32.47 ± 25.964 (95% CI .071 to 27.914), $p = .046$. Like the Guardian, Telegraph and BBC, the visually improved D2 version of Reuters was considered significantly less biased than its diminished D3.

6.9 Results

Table 6-2 also displays the results of the Simple Main Effects for Distortions statistical analysis that were previously presented in section 6.8.2. It depicts nine, One-Way Repeated Measures ANOVA's using the Bonferroni correction for multiple comparisons. Statistically significant results from pairwise comparisons ($p < .050$) are highlighted for the Guardian (dark blue), Telegraph (light orange), BBC (dark orange), and Reuters (dark grey). The colours are the same as those used for each website in the graph shown in Figure 6-3. Other interesting results ($p < .200$) are highlighted in light grey. All tests passed Mauchly's Test of Sphericity, the results of which are displayed above each table. Mean and Standard Error are reported in the corresponding Row and Column of each table. F Statistics and Pairwise P Values for the pairwise comparisons are reported in the intersecting cells.

Table 6-2 includes nine individual tables in the style of distance tables commonly seen on maps. Consequently, the data is inverted on the diagonal to allow for comparisons. The blank cells indicate where a comparison is impossible, e.g. D0 to D0. These tables display the Mean and Standard Error of each distorted webpage/article combination as well as the F statistic and p value results of the pairwise comparisons conducted as part of the Simple Main Effects for Distortions. Above each is the results of the Mauchly's Test of Sphericity (MTS). Significant results are highlighted for the Guardian (dark blue), Telegraph (light orange), BBC (dark orange), and Reuters (dark grey). The colours correspond to Figure 6-3. Other interesting results ($p < .200$) are highlighted in light lighters versions of each colour.

As can be seen by this Gestalt overview, there is a distinct difference in perceived bias in the webpage/article combinations with high quality and low-quality aesthetics. This is further demonstrated with five significant results further emphasising the impact of aesthetics.

Table 6-2 Results of nine, One-Way Repeated Measures ANOVA's using the Bonferroni correction for multiple comparisons.

Guardian – MTS $\chi^2(2) = 3.439, p = .625$					Telegraph – MTS $\chi^2(2) = 3.095, p = .685$					Independent – MTS $\chi^2(2) = 6.282, p = .280$				
	D0	D1	D2	D3		D0	D1	D2	D3		D0	D1	D2	D3
	M 70.51 SE 4.334	M 63.29 SE 4.136	M 53.84 SE 4.707	M 75.16 SE 2.776		M 54.27 SE 4.160	M 44.60 SE 4.482	M 35.733 SE 4.263	M 58.422 SE 4.730		M 26.978 SE 3.601	M 30.400 SE 4.106	M 25.867 SE 3.393	M 36.600 SE 4.844
D0	M 70.51 SE 4.334	F 5.442 p 1.00	F 5.442 p .061	F 5.442 p 1.00	D0	M 54.27 SE 4.160	F 5.856 p .698	F 5.856 p .020	F 5.856 p 1.00	D0	M 26.978 SE 3.601	F 1.325 p 1.00	F 1.325 p 1.00	F 1.325 p .720
D1	M 63.29 SE 4.136	F 5.442 p 1.00	F 5.442 p .660	F 5.442 p .107	D1	M 44.60 SE 4.482	F 5.856 p .614	F 5.856 p .116	F 5.856 p 1.00	D1	M 30.400 SE 4.106	F 1.325 p 1.00	F 1.325 p 1.00	F 1.325 p 1.00
D2	M 53.84 SE 4.707	F 5.442 p .061	F 5.442 p .660	F 5.442 p .004	D2	M 35.733 SE 4.263	F 5.856 p .020	F 5.856 p .614	F 5.856 p .007	D2	M 25.867 SE 3.393	F 1.325 p 1.00	F 1.325 p 1.00	F 1.325 p .493
D3	M 75.16 SE 2.776	F 5.442 p 1.00	F 5.442 p .107	F 5.442 p .004	D3	M 58.422 SE 4.730	F 5.856 p 1.00	F 5.856 p .116	F 5.856 p .007	D3	M 36.600 SE 4.844	F 1.325 p .720	F 1.325 p 1.00	F 1.325 p .493
Economist – MTS $\chi^2(2) = 1.891, p = .864$					Spectator – MTS $\chi^2(2) = 10.497, p = .062$					New Statesman – MTS $\chi^2(2) = 9.114, p = .105$				
	D0	D1	D2	D3		D0	D1	D2	D3		D0	D1	D2	D3
	M 30.27 SE 4.103	M 41.00 SE 26.962	M 35.33 SE 26.795	M 43.07 SE 31.444		M 36.02 SE 26.995	M 41.80 SE 31.119	M 34.69 SE 26.805	M 49.93 SE 27.952		M 20.022 SE 3.799	M 19.467 SE 2.672	M 18.089 SE 2.720	M 24.667 SE 3.884
D0	M 30.27 SE 4.103	F 1.965 p .321	F 1.965 p 1.00	F 1.965 p .237	D0	M 36.02 SE 26.995	F 2.671 p 1.00	F 2.671 p 1.00	F 2.671 p .133	D0	M 20.022 SE 3.799	F .699 p 1.00	F .699 p 1.00	F .699 p 1.00
D1	M 41.00 SE 26.962	F 1.965 p .321	F 1.965 p 1.00	F 1.965 p 1.00	D1	M 41.80 SE 31.119	F 2.671 p 1.00	F 2.671 p 1.00	F 2.671 p .855	D1	M 19.467 SE 2.672	F .699 p 1.00	F .699 p 1.00	F .699 p 1.00
D2	M 35.33 SE 26.795	F 1.965 p 1.00	F 1.965 p 1.00	F 1.965 p 1.00	D2	M 34.69 SE 26.805	F 2.671 p 1.00	F 2.671 p 1.00	F 2.671 p .125	D2	M 18.089 SE 2.720	F .699 p 1.00	F .699 p 1.00	F .699 p 1.00
D3	M 43.07 SE 31.444	F 1.965 p .237	F 1.965 p 1.00	F 1.965 p 1.00	D3	M 49.93 SE 27.952	F 2.671 p .133	F 2.671 p .855	F 2.671 p .125	D3	M 24.667 SE 3.884	F .699 p 1.00	F .699 p 1.00	F .699 p 1.00
Al Jazeera – MTS $\chi^2(2) = 1.130, p = .951$					BBC – MTS $\chi^2(2) = 2.135, p = .830$					Reuters – MTS $\chi^2(2) = 3.966, p = .554$				
	D0	D1	D2	D3		D0	D1	D2	D3		D0	D1	D2	D3
	M 42.778 SE 4.560	M 47.867 SE 4.259	M 38.244 SE 4.299	M 44.400 SE 4.013		M 31.867 SE 4.274	M 26.533 SE 3.923	M 23.244 SE 3.293	M 39.444 SE 3.468		M 22.356 SE 3.730	M 23.489 SE 3.766	M 18.422 SE 2.871	M 32.467 SE 3.870
D0	M 42.778 SE 4.560	F .858 p 1.00	F .858 p 1.00	F .858 p 1.00	D0	M 31.867 SE 4.274	F 3.506 p 1.00	F 3.506 p .580	F 3.506 p 1.00	D0	M 22.356 SE 3.730	F 2.690 p 1.00	F 2.690 p 1.00	F 2.690 p .494
D1	M 47.867 SE 4.259	F .858 p 1.00	F .858 p .805	F .858 p 1.00	D1	M 26.533 SE 3.923	F 3.506 p 1.00	F 3.506 p 1.00	F 3.506 p .097	D1	M 23.489 SE 3.766	F 2.690 p 1.00	F 2.690 p 1.00	F 2.690 p .481
D2	M 38.244 SE 4.299	F .858 p 1.00	F .858 p .805	F .858 p 1.00	D2	M 23.244 SE 3.293	F 3.506 p .580	F 3.506 p 1.00	F 3.506 p .020	D2	M 18.422 SE 2.871	F 2.690 p 1.00	F 2.690 p 1.00	F 2.690 p .046
D3	M 44.400 SE 4.013	F .858 p 1.00	F .858 p 1.00	F .858 p 1.00	D3	M 39.444 SE 3.468	F 3.506 p 1.00	F 3.506 p .097	F 3.506 p .020	D3	M 32.467 SE 3.870	F 2.690 p .494	F 2.690 p .481	F 2.690 p .046

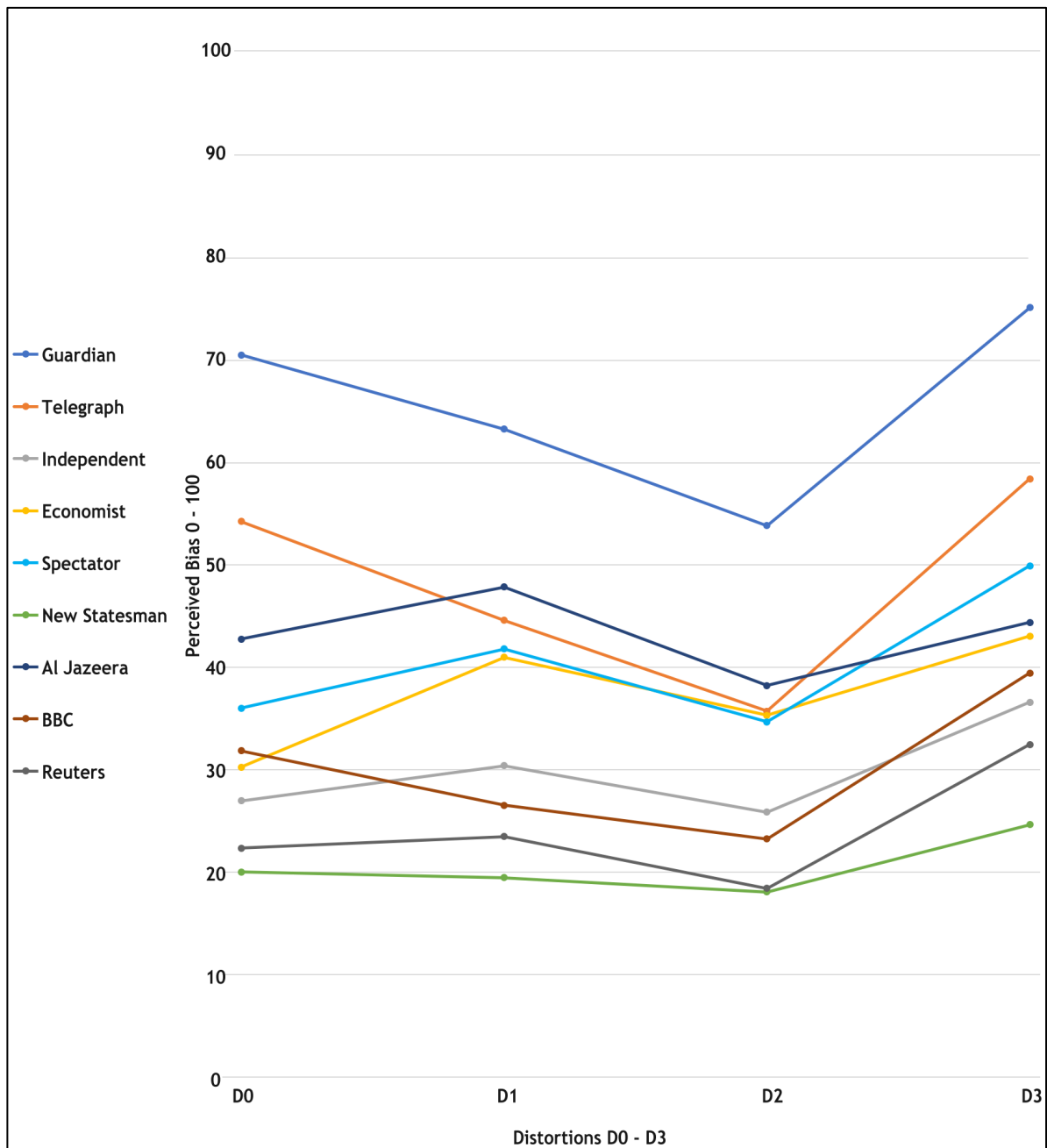


Figure 6-3 The impact of increasing and decreasing the visual quality of the aesthetics for each webpage/article combination on perceived bias.

6.10 Findings from the Experiment Results

6.10.1 Comparing the D1 Controls to their Respective D2 High Quality Aesthetics

The D1 controls for the nine webpage/article combinations can be compared to their respective D2 aesthetically high quality distorted webpage/article combinations in the results of the statistical analysis shown in Table 6-2, and depicted in Figure 6-3. These are based on the results of the statistical analysis in

section 6.8.2. For each webpage/article combination, the increase in the quality of the aesthetics, decreased the perception of bias. The effect is common across the three categories of news websites and across articles with different levels of bias in the content. Like experiment two, there is a noticeable difference in the level of effect depending on how biased the news article is. There was a greater reduction in the perception of bias for those articles with higher levels of bias, while there was less of an effect for those with a lower level of bias. However, despite some evident trends in the data, there are no significant differences in the perception of bias between any of the D1 and their respective D2 distorted webpage/article combinations.

6.10.2 Comparing the D1 Controls to their Respective D3 Low Quality Aesthetics

Comparing the D1 controls for each of the nine webpage/article combinations to their respective D2 aesthetically low-quality distorted webpage/article combinations, a clear and common trend is visible in the data. When the quality of the aesthetics is reduced, the perception of bias is increased. This effect was common across all three categories of news website and across articles with different levels of bias in them. While no significant differences were found between any of the D1s and their respective D2 distortions, there were several results that were close. These are shown in light grey in Table 6-2.

It should be noted that the level of effect of the D3 distortion in increasing the perception of bias was much stronger than the D2 for decreasing the perception of bias. This is due to the relative ease in decreasing the quality of the aesthetics of a webpage than to increase the quality of the aesthetics of an already high-quality professional news agency's webpage.

6.10.3 Comparing the D2 High Quality Aesthetics to their Respective D3 Low Quality Aesthetics

From the results of the statistical analysis in section 6.8.2, the results presented in Table 6-2, and depicted in Figure 6-3, there is a clear and obvious difference in the perception of bias between the aesthetically high quality D2 webpage/article combinations and their respective aesthetically low quality D3 webpage/article combinations. This effect was consistent across all three categories of news websites and across articles with different levels of bias. This is backed up by significant differences between the distortions on four of the websites: Guardian, Telegraph, BBC, and Reuters. Based on these results it is possible to say that perceived bias within the same article could be considered significantly more or less biased based on the quality of aesthetics of the news webpage it is presented in.

While the effects of aesthetic quality may not be as obvious when one compares the control D1 to either their respective high quality D2s or low quality D3s distortions, it is extremely obvious when they are compared to each other. This is important as news articles are written on the same topics daily and presented in different ways.

6.11 Hypotheses Results

H₀ Perceived bias will not be increased due to a reduction in the quality of visual presentation.

H_A Perceived bias will be increased due to a reduction in the quality of visual presentation.

The answer to the first hypothesis can be determined by inspecting Table 6-2 and Figure 6-3 to ascertain if perceived bias was higher in the D3 aesthetically low quality webpage/article combinations compared to their respective D1 control webpage/article combinations. As is clearly shown, eight of the nine examples showed some increase.

While no results proved significant, three of the webpage/article combinations, the Guardian $p = .107$, the Telegraph, $p = .116$, and the BBC $p = .097$ showed marked increases. Only Al Jazeera demonstrated a decrease in perceived bias. It should be noted however, that Al Jazeera demonstrated an increase in perceived bias when its D3 is compared to its D2 showing that the effect was at least partially in place when compared to the improved version of the website. However, despite the pattern shown in these results, the null hypothesis cannot be rejected.

It should be noted, that despite not rejecting the null hypothesis due to no significant differences, that there does appear to be some correlation between reduction in visual quality in a news webpage and an increase in perceived bias in a news article contained therein. It is also likely that if the visual quality of each webpage/article combination was further reduced that there would have been some significant differences found. Thus, this warrants further investigation and is discussed further in the Discussion section 6.13.

H₀ Perceived bias will not be decreased due to an improvement in the quality of visual presentation.

H_A Perceived bias will be decreased due to an improvement in the quality of visual presentation.

The second hypothesis can be answered by comparing the D2 aesthetically high quality webpage/article combinations to their respective D1 control webpage/article combinations in Table 6-2 and Figure 6-3. While none of the results of the pairwise comparisons of the individual One-Way Repeated Measures ANOVAs were significant, the trend depicted in Figure 6-3 is obvious. Every webpage/article combination showed a distinct decrease in perceived bias when the aesthetics was distorted to improve the visual presentation of the news. It should be noted that improving the visual quality of the aesthetics of the webpages of international news organisations who spend millions on their websites is a much more difficult task than reducing the visual quality of the webpages. It should also be noted that it is virtually impossible to create a news article which a large population of users will consider unbiased.

The lower the rate of perceived bias in each article, as rated in the D1 control webpage/article combinations, the lower the decrease in perceived bias in D2. That is, there is only scope for a marginal reduction in bias when there is less bias in the original article. This is most evident in the webpage from the New Statesman, which was paired with the least biased article, as rated by participants in the pre-test experiment and the

full experiment and showed the smallest decrease. This indicates that there is a minimum baseline or floor which even unbiased articles cannot descend beneath. This article also had a large amount of data in the text which may indicate participants relied on some form of numbers heuristic when evaluating it. Previous research has shown that users may find information more credible and persuasive with the inclusion of numbers (Yalch & Elmore-Yalch, 1984). It is possible that a similar effect may impact perceived bias, which is a core dimension and measure of credibility, especially when judging news online. Conversely, the greater the perceived bias in the D1 version of each webpage/article combination, e.g. Guardian, Telegraph and Al Jazeera, the greater the reduction in perceived bias in their respective D2s.

However, despite this definite and common decrease in perceived bias when the D2s are compared to their D1, from D1 to D2 as demonstrated in Table 6-2 and Figure 6-3 the null hypothesis cannot be rejected as none of the results proved significant.

H₀ There will be no difference in the perception of bias between high-quality and low-quality versions of the same news article.

H_A There will be a difference in the perception of bias between high-quality and low-quality versions of the same news article.

A comparison of the results in Table 6-2 and Figure 6-3 of the aesthetically high quality D2 webpage/article combinations to the aesthetically low quality D3 webpage/article combinations for each website provides the answer to the third hypothesis. The results show a definite increase in perceived bias. This effect is common across all webpage/article combinations and across each category of news website and across articles containing different levels of bias. More importantly an inspection of Table 6-2 shows four significant differences between the perceived bias in their D2 and D3 distortions. Guardian $p = .004$, Telegraph $p = .007$, BBC $p = .020$, and Reuters $p = .046$. As stated already, it must be noted that it was difficult to improve upon the quality of the D2 webpage/article combinations. At the same time, it would be easily possible to further degrade the quality of the visual experience in D3. Based on this, the Null hypothesis can be rejected and the alternative accepted.

6.12 Design Implications

This experiment demonstrated that it is possible to predicatively increase and decrease the perception of bias in a news article by decreasing or increasing the quality of the aesthetics of the news webpages that contain them. This effect was common across the majority of websites and across news articles with a range of biased content.

There was one overarching theme behind the individual distortions which caused this effect. That theme was the mirroring of the aesthetics of news websites which attempt to generate as much revenue as possible through their websites.

The implications for news website designers are simple. A reduction in the quality of the aesthetics will increase the perception of bias in a news article, conversely an increase in the quality of the aesthetics will decrease the perception of bias in the article. Using this research as a base, news website designers should attempt to identify elements of their news website design which may be impacting perceived quality. This could be through their presence or through issues with their implementation.

6.13 Discussion

This experiment demonstrates that perceived bias in a news article, can be decreased or increased through improving or diminishing the visual quality of the aesthetics. The results showed significant differences in perceived bias between the high-quality D2 and the low-quality D3 versions of the same webpage/article combinations, see Table 6-2 and Figure 6-3. Furthermore, while there are no significant differences when comparing the D2s and D3s to their respective D1 control distortions, the trends are evident in Figure 6-3. When the visual quality of the aesthetics of a news webpage are decreased, perceived bias in the news article increases. Conversely, when the visual quality of the aesthetics of a news webpage increases, perceived bias in the news article it contains decreases. It is possible that participants were judging perceived bias by some form of unknown heuristics, though further research is required to identify their nature.

It is likely that the trends evident in Table 6-2 and Figure 6-3 would have been more pronounced if the visual quality of the of the D2s could have been further increased and the visual quality of the D3s further decreased. Although it would be difficult to improve upon the visual aesthetic of the webpages of such internationally successful news organisations while keeping their functionality intact, it is possible that an alternative distortion process than that described in Table 6-1 might have resulted in aesthetically higher quality D2s. Conversely, it would be relatively easy by comparison to further reduce the aesthetic quality of the D3s. This would likely have resulted in some additional significant results which may have made it possible to reject the first two null hypothesis in section 6.11.

Participants probably associated low-quality visual presentations as being more likely to be purveyors of biased news. This is a serious issue for news agencies. The distortion techniques, such as the addition of extremely low-quality advertisements and large calls to action, adopted in this experiment to decrease the quality of the visual presentation are all in common usage on content farms, arguably the lowest tier of online news 'organisations'. It may only be a matter of time before they are seen in more mainstream news websites.

6.14 Lessons Learned from Experiment Three

Overall, experiment three worked well. This was likely due to the maturing of the CAFE experiment framework and experiment platform. There are two main lessons from experiment three.

- One element which may have impacted the experiment was the content of the advertising. While all the advertisements were selected from existing low quality content farms and/or serious news websites for each of the distortions, the selection process did not take into account the variability of the content of such advertisements.

Due to cross website tracking and user profiling, the advertising displayed to different users varies a lot. While every attempt was made to ensure that the type, style, and content of the advertisements included in this experiment was representative of the type many users experience when visiting the different types of news websites, it is possible that different advertisements might have affected the experiment results differently. This should be considered for future work in the domain.

- One fact evident from the results is the difficulty of improving the quality of the aesthetics of the webpages from the nine websites chosen for this experiment. These are some of the most visited news websites in their respective markets and each spends a great deal of money to maintain high quality aesthetics on each webpage.

In hindsight, webpages from existing low quality content farm news websites should have been included in the experiment. Thus, it would then have been possible to attempt a crossover experiment. Whereby webpages from each type of news website would be distorted to look like webpages from websites at the other end of the news website spectrum. Not only would this have ensured that the experiment more closely reflected reality and acted as a double confirmation of the results, it is also likely that the differences in perceived bias in the news articles contained in such polar-opposite webpages would have been judged as significantly different. Thus, further proving the supposition behind this research.

6.15 Summary of Experiment Three Findings

The main finding from this experiment is that perceived bias in a news article can be predictively impacted by distorting the quality of the aesthetics of a news webpage. This also provides further confirmation of the importance and influence of website design characteristics such as professionalism which was investigated in experiment two. It also demonstrates the influence such characteristics of a news website's

design have on the perception of the message contained in the content of individual news articles. This finding beseechs the questions, which characteristics influence which dimension of credibility, to what degree, and can this be harnessed?

This finding has practical importance. Many news agencies that are following a low-cost content and online advertising business model are engaged in a race to the bottom which is unsustainable. Like experiment two, this research also has practical concerns for news consumers. It has demonstrated that unbiased news online could be considered more biased by presenting it in webpages with low quality aesthetics. Conversely, and more worryingly, the opposite is also true. Biased news could be considered less biased by presenting it in news webpages with high quality aesthetics.

6.16 Conclusions

The importance of news to society cannot be overemphasized. This experiment demonstrated that perceived bias within the same news articles can be manipulated through distorting the visual quality of its presentation. Aesthetically low-quality webpages increased perceived bias in the article contained therein, while aesthetically high-quality decreased perceived bias. Significant differences were found between the low quality and high-quality distortions of the same webpage/article combination.

“The only thing necessary for the triumph of evil is for good men to do nothing.” Edmund Burke

7 Summary Findings

7.1 *State of the Art*

The State of the Art presented in chapter 2 satisfied the first research objective, namely: To research the State of the Art on measuring bias and identify the frameworks and processes which have been used. It also delivered a minor contribution, namely undertaking an in-depth review of existing frameworks and the underlying theory in the domain. There are five main findings from the State of the Art detailed below.

7.1.1 *Isolated Research Efforts*

The main finding from this is the lack of standardisation and the disjointed nature of the domain e.g. studies with inconstant and overlapping definitions, which suggests that much of the work was undertaken in isolation, or at least without a sense of the overall domain. There are two means by which this is apparent. First, no example was found of a study being repeated by the same researchers or by others. Second, many studies also fail to adequately cite relevant research. Third, by a lack of agreed definitions, terminology, and overlapping or dual classifications of bias.

7.1.2 *A Need for a Classification of Biases that Impact the Different Stages of News*

Whether cause or effect, the lack of coordinated research whereby delineation, categorisation, and definitions for each type of bias are settled is non-existent. Even though researchers regularly acknowledge the lack of such theoretical underpinnings has failed to incite efforts to address this lacuna. Consequently, an initial classification of the biases which affect the production stage of news was put forth in Table 9-2, Table 9-3, Table 9-4, and Table 9-5. An initial classification was also put forth for the main types of dissemination biases that impact print, radio, television, and the WWW. These can be seen in Table 9-6, Table 9-7, Table 9-8, and Table 2-3. This classification will also help to solve the issue of research being conducted in isolation. This classification of biases should also help with the issue of overlapping, contradicting, or the complete lack of definition for many of the different forms of news bias.

7.1.3 *Bias in the Dissemination of News on the WWW has yet to Receive the Attention it has in Traditional Mediums*

This disparity is most evident in the gestalt overview provided by comparing the classification of biases affecting traditional mediums which can be seen in Table 9-6, Table 9-7, and Table 9-8, to the lack of studies on bias in the dissemination of news on the WWW shown in Table 2-3.

7.1.4 Most Studies on Bias in the News do not Focus on the Content or the Message Therein

A surprising finding is the fact that most studies on bias in the news do not focus on the content or the message contained therein. Instead most studies focus on surrounding factors such as the facial expressions of news readers (H. S. Friedman et al., 1980; Miller et al., 2007; Zimmerman, 2013), or the content or favourability of images (Hehman et al., 2012; Towner, 2017). There are three related reasons for this. First, because bias is most often found in emotive or controversial topics, it is possible that our own biases may act as confounding variables in any study of the content. Second, because it is easier to measure the amount of column inches or the favourability or warmth of facial expressions of news anchors than it is to measure bias in text. Third, because bias is subjective, studies which focus on the content are less likely to find clear patterns. This knowledge may help future researchers to decide to focus on the difficult task of identifying and measuring bias in the content and message of an article, radio show, or television show.

7.1.5 The Lack of Theoretical or Empirical Models or Frameworks to Explain How Judgements of Bias in News are Arrived at and their Effect, or how to Measure it.

There is a lack of models, theories, or frameworks to explain how humans make judgements of bias online, or in any other medium. This was a surprising finding considering the long history of research and the divisive nature of the construct. The lack of underlying theory is major gap in the body of knowledge and one which will hopefully be addressed in the future. This search also demonstrated that there are no empirical frameworks or semi-standardised methods or techniques for measuring bias online or in any other domain. Remedying this would be a worthwhile for future researchers in the domain in the years to come.

7.2 Credibility Assessment Framework for Experimentation (CAFE)

The need for the design and development of the Credibility Assessment Framework for Experimentation (CAFE) and accompanying opensource experiment platform was identified in the State of the Art review. It satisfies the second research objective. Namely, the design, development, and testing of an innovative framework and platform that supports a range of experimental designs to detect and record user's perception of bias. It also delivered the second minor research contribution, the CAFE experiment framework and experiment platform.

7.2.1 An Adaptable Framework and Opensource Platform to aid in the Design, Development, and Deployment of Experiments to Measure Credibility or its Dimensions and Measures

The open framework is designed to be adapted by the user to suit their experiment and research aims. There are several benefits to using CAFE, including:

- The primary benefit is its ability to help researchers identify issues or conflicts in their experiment design early. This can save significant time and help prevent critical issues with an experiment satisfying its research aims.
- It is designed to formulate the researchers' thoughts and to work as both a plan and as a record of the experiment process.
- It helps to make sure that the understanding of all the members of a team is in unison.
- It can be used as an aid to help explain or propose an experiment to a supervisor, principle investigator, review panel, ethics committee, or other interested party.
- It acts as a record of decisions.
- It can help with working out experiment requirements and timelines.

An accompanying experiment platform was also developed and made opensource so that the research community can quickly deploy experiments. This has been used in several experiments, both as part of this research and by others, and is in ongoing use. As per section 3.8, version two will include facilities designed to make it easier to deploy by those without coding experience and to increase its functionality. Additional functionality will be added to this in future iterations. This includes administrator control panels to reduce the amount of code editing for novice or non-technical researchers and the ability to deploy additional types of experiments.

7.2.2 Eight Existing Best Practices, Recommendations, and Novel Contributions for Experimentation

The CAFE framework also highlighted eight existing best practices, recommendations and novel contributions for eight different stages of a typical experiment to measure credibility online. While many of these such as Statistical Power Analysis and crowdsourcing have been recommended best practices or standard techniques in the overarching domains of behavioural psychology or computer science, they have yet to be widely adopted in credibility or bias research. Some, such as the use of VAS as a new form of measurement scale, have until now not been used in either domain. The repository and classification of credibility measures is also a significant development in the domain, see section 3.6.3.

7.3 Experiment One Findings

There were three main findings from experiment one, reported in chapter 4; the experiment results, the results of the reflective questions, and the demonstration of the CAFE framework and accompanying platform. Experiment one partially satisfied the third research objective; To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impact perceived bias. It also partially satisfied the major research contribution; The identification and validation through

experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

7.3.1 The Impact of Individual Webpage Features, User Characteristics, and News Website Category on the Perception of Bias

The experiment had three main threads of investigation. First, individual features of a news webpage proved inconsistent in their influence on the perception of bias. However, based on the underlying trends evident in the data, it is likely that if the visual impact of the distortions was increased that the impact would have been more consistent and predictable. Consequently, the main finding from this is that it should be investigated further in future. Second, while some user characteristics proved to be predictive influencers of perceived bias, the lack of consistency of the results means that further investigation is also required to understand them better. Third, the experiment results showed that the category of news website had a definite impact on the perception of bias, with the webpage/article combinations from traditional print news websites being perceived as less biased than those of international multi-format news agencies or news magazines. However, these trends were insufficient to categorically reject the null hypothesis. This should also be investigated further in future.

7.3.2 Reflective Questions Reveal the Characteristics of a Website's Design or Webpage's Aesthetics that Impact Perceived Bias the Most

The most important findings from the reflective questions section of the first experiment were the four characteristics which participants believed could reduce the perception of bias. These were; Seriousness, Professional, High Quality Aesthetics, and Good Design.

7.3.3 Initial Demonstration and Testing of the CAFE Framework and Experiment Platform

The third major findings from the first experiment was an initial demonstration of the CAFE framework and the accompanying experiment platform. The lessons learned from implementing and executing the first experiment were integrated into subsequent versions of both. It also demonstrated the success of the overall experiment process.

7.4 Experiment Two Findings

Experiment two had four main findings; the experiment results, the novel pre-test content preparation techniques, increased understanding of bias for a proposed future model, and further demonstration and evolution of the CAFE framework and platform. Experiment two partially satisfied the third research objective; To identify whether features and characteristics of a news website's design, reflected in a news webpage's aesthetic, impact perceived bias. It also partially satisfied the major research contribution; The

identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

7.4.1 Demonstration of the Impact of Characteristics of Website Design Reflected in a Webpage's Aesthetics Impact Perceived Bias

The main finding from experiment two is the knowledge that characteristics of a news website's design reflected in a news webpage's aesthetic can greatly impact the perception of bias within the news articles they contain. This finding was largely consistent across different categories of news websites and across news articles ranging from relatively unbiased to extremely biased. These findings are important to both news website designers and editors, and news consumers. It is the first time that the visual presentation of information on news websites has been shown to consistently impact the perception of bias in the articles they contain. As the perception of bias in a news article increases, the perception of credibility of the news organisation behind it decreases (Fico et al., 2004).

7.4.2 Increased Understanding of Bias for a Proposed Model to Explain How Judgements of Bias are Made and their Effect

Like experiment one, experiment two also included a number of reflective questions on a range of issues relating to news consumption, its visual presentation, and bias. The increased understanding of how users form judgements of bias online will contribute to the development of a model or models in future.

7.5 Experiment Three Findings

Although the first two null hypotheses in experiment three were not rejected, the rejection of the third hypothesis resulted in two main findings; confirmation that characteristics of a news website's design reflected in a webpage's aesthetic, impact the perception of bias in the articles they contain; and demonstration that bias can be predictively influenced with such. Experiment three satisfied the fourth research objective by focusing on characteristics; To evaluate the adaptation of characteristics of a website's design, reflected in a news webpage's aesthetic, to predict the perception of bias. It also validated the major research contribution, the identification and validation through experimentation of the impact of features and characteristics of news website's design, reflected in a news webpage's aesthetic, on users' perception of bias.

7.5.1 Confirmation that Characteristics of a Website's Design, Reflected in a Webpage's Aesthetics, Influence Perceived Bias in the News Article Contained Therein

The first finding from experiment three confirms the results of the second experiment, namely that individual characteristics a news website's design and/or of a news webpage's aesthetics can impact perceived bias in the news articles contained therein. In this instance, the characteristic of aesthetic quality,

which was also identified in the reflective questions of experiment one, was used as the influencing variable. While closely aligned to the characteristic of professionalism which was used in experiment two, aesthetic quality is different in that it focuses on the quality of the overall visual aesthetic which is impacted by the quality and balance of the combination of elements. Whereas the characteristic of professionalism refers to how well the features and elements that make up a news webpage were implemented.

7.5.2 Demonstration that Perceived Bias can be Predictively Influenced

The second major finding from experiment two is that perceived bias can be predictively influenced. The experiment found that increasing the visual quality of a news webpage's aesthetic decreased the perception of bias while decreasing the visual quality increased the perception of bias. The effects were common across all the websites tested and across news articles with different levels of bias.

“It is easier to fool a fool than to convince a fool they have been fooled” - Unknown, but commonly attributed to Mark Twain

8 Conclusions

Increasingly, consumers are turning to the WWW as their primary medium of news consumption. The impact of its emerging dominance over traditional mediums is compounded as more of the world’s population gain access to it. Despite this, not enough is known about the effects of medium on the dissemination of news. This research has shown that the impact of news websites on the news articles they contain is especially worrisome. Specifically, it has shown that a consumer’s perception of bias within a news article can be affected by its visual presentation, and what is more, that it can be predictively influenced. Bias is a constant concern and a common accusation of news consumers. It also has a detrimental effect on a news organisation’s credibility.

This thesis began with a comprehensive State of the Art review, which includes the largest classification of definitions of bias existent in the literature. It also includes the first classification of biases which can influence the production of news, and the first classification of biases that can impact the dissemination of news across the four main mediums, print, radio, television, and the WWW. The State of the Art also highlights some of the most prominent research in the domain, lamenting the lack of underlying theory. It also demonstrated the lack of empirical frameworks for conducting research into bias. Consequently, the overarching domain of credibility was explored. Ten human-centric MTFS were identified which explain how judgements of credibility are made online. The underpinning theory for the majority of the MTFS, namely the ELM, HSM, and the C&AHIP were also explored and a visual realisation of the HSM was also created. The lack of an empirical framework for measuring credibility provided an opportunity to develop the CAFE framework and accompanying experiment platform, which can also be used to measure any of the dimensions and measures of credibility, such as bias.

This thesis confirmed earlier work by Fogg proving that bias is a core dimensions and measure of credibility when judging news online (Fogg et al., 2003). It has shown that perceived bias, like perceived credibility, is affected by the visual presentation of a news webpage, by demonstrating that individual features and characteristics of a news website’s design, reflected in a news webpage’s aesthetic, can impact the perception of bias. It showed that as the level of professionalism of a webpage’s aesthetic decreases, the perception of bias increases, and vice versa. Most importantly, it also revealed that perceived bias in a news article can be predicatively influenced by adapting visual characteristics of the webpage’s aesthetic such as its visual quality. This knowledge is important as it opens up new directions for future research.

Bias of one sort or another is endemic in news. It is a complex and multifaceted construct, yet at its root, it can simply be described as imbalance or slant. Nevertheless, its effect can be severe. It may be deliberate

or accidental, known or unknown, real or perceived. It may affect any stage of the news pipeline; production, dissemination, or consumption. Most academics and journalists agree that it is nigh on impossible to publish news that is completely unbiased, and consumers should not expect such. Instead, they should arm themselves with the knowledge and ability to detect bias and neutralise its effect. The easiest way of combating bias is to consume news from multiple sources with opposing points of view. Future work in the domain should look at this more closely.

Most news consumers have at least a rudimentary awareness of bias in the news. A rudimentary understanding includes the knowledge that journalists may bias reporting with the addition of personal opinion, or by imbalanced presentation of arguments and facts. More knowledgeable consumers may be able to define and recognise different types of bias, such as framing, agenda setting, or selection bias. While those with a deep understanding may even be able to identify the more difficult or hidden forms such as, gatekeeping, intergroup, stereotype, or description bias. This awareness means that there is at least a certain amount of oversight or scrutability of the news by researchers and the end consumer. However, until now, the impact of the visual presentation of a news website or webpage on the perception of bias in the news articles they contain was unknown. This research will hopefully help consumers to realise that the websites and webpages surrounding and delivering news articles could be impacting their perception of bias and other dimensions and measures of credibility.

8.1 Research Objectives and Achievements

The research question answered by this thesis is: “Can the features and characteristics of a news website’s design, reflected in a news webpage’s aesthetic, influence users’ perception of bias, and if so, to what extent can their deliberate use reasonably predict that perception of bias?” To answer this question, four research objectives were derived. Fulfilling these objectives provides one major and two minor research contributions, discussed in section 8.2, below.

8.1.1 Objective One: Research the State of the Art on Measuring Bias and what Frameworks are Used

To address the first research objective a comprehensive review of the domain was undertaken. The State of the Art began with an introduction to the overarching research area of credibility, see section 2.2.1. This was followed by four categories of definitions of the construct are provided in section 2.2.4. The review of the domain quickly identified ten, human-centric models, theories, frameworks, and schematics (MTFS) to explain how judgements of credibility are made online. These are presented in section 2.2.8 and contrasted in Table 2-1. The majority of these rely on underlying theory provided by Petty and Cacioppo’s Elaboration Likelihood Model (ELM), Chaiken’s Heuristic-Systematic Information Processing Model (HSM), and Shiffrin and Schneider’s general theory resulting from their Controlled and Automatic Human

Information Processes (C&AHIP), which are detailed in section 2.2.7 (Chaiken, 1980; Petty & Cacioppo, 1986; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). As part of this review, the first realisation of the highly cited HSM was also provided in Figure 2-4. Sections 0 connects these underlying theories to the MTFs while section 2.2.11 highlights an extensive compendium of research demonstrating the impact of features and characteristics of a website's design, reflected in a news webpage's aesthetic, on the perception of credibility. Issues with current approaches to credibility experimentation online and the benefits of an empirical framework and accompanying experiment platform are then discussed in sections 2.2.12 and 2.2.13. As no empirical frameworks were discovered, it was decided to create an empirical framework and experiment platform that would be capable of measuring credibility or any of its individual dimensions and measures.

The State of the Art then focused on the specific research area of bias. This began with an introduction to the construct in section 2.3.1. This was followed by a section highlighting the lack of underlying theoretical research on the topic, 2.3.2. A first of type classification of definitions of bias from the literature is then provided in section 2.3.3. This serves to demonstrate the wide verity of interpretations of the construct. Making this available to other researchers in the domain is a core aim of future work. Breaking down the types of definitions may well be a first step in understanding the dimensions of news bias. The chapter then provided a classification of biases which affect the production (section 2.3.4.1) and dissemination (section 2.3.4.2) of news across all four mediums. The lack of research on bias in the consumption of news on the WWW is clearly shown by comparing the depth and range of studies on bias on the radio (Table 9-6), print (Table 9-7), and television (Table 9-8), to the WWW (Table 2-3). This will hopefully help focus future research on the WWW which has become the main medium for accessing news for many. Though out of scope for this research, an initial classification of biases which impact the consumption of news was also provided, see section 2.3.4.3. Together, these will form the backbone of a major effort of future work to provide a comprehensive classification and definitions of biases, that impact the production, dissemination, and consumption of news. Details of an initial classification, including the difficulties of creating it, have recently been made available (Spillane & Wade, 2020). An in-depth review of research was then undertaken of bias in the dissemination of news in traditional mediums (section 2.3.5) and on the WWW (section 2.3.6). This in-depth review revealed no theoretical or empirical frameworks to provide understanding of the construct or to guide research into the construct in any medium. This is further discussed in sections 2.3.7 and 2.3.8.

To demonstrate the lack of theoretical and empirical frameworks for bias, the aligned domain of trust was then investigated, see section 2.4. Trust, like bias, is a perceived value that individuals put on information or a person when making judgements about its acceptability, see section 9.8.1. Both are core dimensions and measures of the overarching construct of credibility. Trust is a very large domain and only a small proportion of it was explored. To begin, definitions of the construct were provided, before studies

demonstrating the impact of website design on trust were reviewed, see section 9.8.2. The review then quickly identified ten theoretical models, and five theoretical frameworks to explain how users form judgements of trust online which are presented in section 9.8.3. As pointed out in the thesis, this disparity is surprising considering the amount of attention and accusations of bias in the news. Surely a first step in identifying whether there is bias is understanding how users make judgements of bias. Section 9.8.4 then presents six empirical frameworks for measuring trust online. Again, these show the disparity in research efforts between bias and trust. While the theoretical models may prove useful in future as a basis for developing a theory for judgements of bias, the fundamental differences in the constructs meant that the frameworks could not be used for measuring the construct. Therefore, the overarching domain of credibility was explored in section 2.2.

8.1.2 Objective Two: To Design, Develop, and Test an Innovative Framework and Platform that Supports a Range of Experimental Designs to Detect and Record User's Perception of Bias

Chapter 3 presents the Credibility Assessment Framework for Experimentation (CAFE). It is an open framework and accompanying experiment platform to aid researchers in conducting experiments to measure credibility, or any of its individual dimensions and measures, such as trust, expertise, or bias. It was designed to simplify the process of designing, developing, and deploying such experiments online. The framework and a detailed description of how it should be used are presented in Figure 3-1 and section 3.3.2. The platform is detailed in section 3.5 including a link to the GitHub repository for the codebase and documentation. It has gone through multiple iterations and is in ongoing further development. It is primarily aimed at novice and non-technical researchers interested in deploying experiments to measure credibility or any of its individual dimensions and measures. One of the core elements of the CAFE framework are the eight existing best practices and novel contributions which it recommends and supports to improve experiments designed to measure credibility or any of its individual dimensions and measures such as bias. These include; the first recommendations for Statistical Power and Effect Size (3.6.1), the advantages of crowdsourcing (3.6.2), and a repository and classification of measures used in peer reviewed academic research to measure credibility online (3.6.3). This repository is the first of its type and will prove a boon to the research community. Also included is a recommendation to use Visual VAS (3.6.4), the benefits of using Latin squares and Latin cubes as experiment matrix (3.6.5), a proposal for weighting individual dimensions and measures to create credibility scores (3.6.6), choosing appropriate statistical tests and detailed reporting assumptions, Confidence Intervals and magnitudes of effect sizes (3.6.7), and lastly, increased reporting of individual measures including the use of visual aids (3.6.8).

There are several benefits to using CAFE for the researcher which are further described in section 3.4. They include:

- It will help researchers to identify potential pitfalls, conflicts, or issues during the planning stage when they can be easily addressed before they reduce the quality or effectiveness of an experiment.
- It formulates the researcher's thoughts. As a realisation of the experiment plan, it forces the researcher to consider each step in the process and to continually revise.
- It helps to make sure that all the members of a research teams understanding is in unison, thus helping to avoid miss communication of research strategy.
- It can be used as an aid to help explain or propose an experiment to a supervisor, principle investigator, or other interested party such as a research ethics committee.
- It acts as a record of decisions and can be easily referenced.
- It can help with working out experiment requirements such as the amount, type and availability of content, while at the same time helping researchers estimate timelines and dependencies for completion.

The CAFE framework and platform went through multiple iterations during its development and use in the design and deployment of the three experiments conducted as part of this thesis. The CAFE framework has also been used as the experimentation platform for two Master's dissertations and one final year research project on bias and credibility, see section 3.3. The next iteration of the platform will focus on updating the platform's code base to the latest standards and to add administrator control panels to reduce the amount of code editing for novice or non-technical researchers and to increase its functionality. The CAFE framework and accompanying experiment platform⁴⁰ satisfy the secondarily research objective.

8.1.3 Objective Three: To Identify Whether Features and Characteristics of a Website's Design, Reflected in a news Webpage's Aesthetic, Impact Perceived Bias

The third research objective was satisfied by the findings of experiments one and two conducted as part of this thesis. The statistical analysis for experiment one is presented in section 4.10. While experiment one found that certain features could impact the perception of bias, the results were inconsistent. The same was true for user characteristics as predictors for perceived bias, see section 4.12. However, different types and categories of news websites showed more consistency, see section 4.11. Through detailed reflective questions, experiment one also identified certain characteristics of a news website's design or a news webpage's aesthetic which participants believe reduce their perception of bias, namely: Seriousness, Professionalism, High Quality Aesthetics, and Good Design, see section 5.2. The design or aesthetic characteristic of professionalism became the focus of the second experiment for two reasons. First, because it has been shown to be a factor of credibility on its own, thus there is a body of literature on it, and second, because it was closely related to seriousness.

⁴⁰ Version 1.0 of the experiment platform may be found at: https://github.com/brendantcd/Credibility_experiment_apparatus.git

Experiment two demonstrated that as the level of professionalism in a news webpage's aesthetic decreased, the level of perceived bias in the news article they contain increased. Conversely, the opposite is also true, when the level of professionalism in a news webpage increased, the level of professionalism decreased. The detailed statistical analysis for this is presented in section 5.9, and the results are presented in Table 5-5, Figure 5-5, and Figure 5-6. The overall findings are presented in section 5.11. Though there were some slight differences, see section 5.11.3, these effects were broadly constant across webpages displaying articles with different amounts of bias.

Experiments one and two, satisfied the third research objective. While features of a news website's design, reflected in a news webpage's aesthetic, are inconstant influencers of perceived bias, characteristics of a website's design or a webpage's aesthetic such as professionalism are consistent across all categories of websites tested and across articles from relatively unbiased to extremely biased.

8.1.4 Objective Four: And if so, To Evaluate the Adaption of Such to Predictively Impact the Perception of Bias

Objective four was satisfied by experiment three. The results of the statistical analysis are shown in section 6.9, the findings are shown in section 6.10, and the hypotheses are discussed in section 6.11. The experiment focused on the quality of the aesthetics and the design of a website, which was the other grouping of characteristics identified by participants in experiment one, see section 4.14.3.3. The experiment was based on the supposition that increasing the quality of the aesthetics would decrease the perception of bias and vice versa. This was based on the knowledge that participants identified this characteristic as possibly reducing the perception of bias. Based on this knowledge, website quality was therefore chosen as an independent variable to decrease and increase the visual quality of the nine webpage/article combinations. The results showed that perceived bias in the news article could be predictively influenced. This effect was common across the majority of websites tested and across articles with different levels of bias, see Table 6-2 and Figure 6-3 in section 6.9. Significant differences were also found in perceived bias in the same news articles displayed in high-quality and low-quality versions of the same news webpage 6.10.3. The experiment also found that it was much easier to reduce the quality of the aesthetics to decrease the perception of bias, than it was to increase the quality of the aesthetics to decrease the perception of bias. The design implications are presented in section 6.12 and the overall results are discussed in section 6.13.

8.2 Contributions to the State of the Art

This thesis adds one major and two minor contributions to the State of the Art. They are as follows:

8.2.1 Major Contribution

The major contribution of this research is *the identification and validation through experimentation of the impact of features and characteristics of a news website's design, reflected in a news webpage's aesthetic, that influence users' perception of bias.*

Characteristics of a news website's design, reflected in a webpage's aesthetic, proved consistent influencers of perceived bias. This research demonstrated that the characteristic of professionalism, which is closely associated with the source, sponsor or the technical team or individual behind a website's implementation, can influence the perception of bias. Perceived professionalism in the design or aesthetic is characterised by how well a design or aesthetic is implemented. This includes issues with misalignment, obviously broken or badly implemented code, or low quality or badly framed or stretched images. As the level of professionalism in the design or aesthetic decreased, the perception of bias increased and vice versa. This effect was common across all nine websites tested and across articles with different levels of bias.

The characteristic of aesthetic quality also proved to be a consistent influencer of perceived bias in a news article. Based on the knowledge that users had identified it in experiment one as being likely to reduce the perception of bias, it was used as a predictive factor in the final experiment. Aesthetic quality is personified by the combination of elements, balance in the design and alignment and positing of elements to create a high quality or low quality experience. By reducing the quality of a website's design or a webpage's aesthetic, it was shown that perceived bias within a news article could be increased. This effect was common across all nine websites tested and across news articles with different levels of bias. The opposite effect was also observed, though it was more muted. When the quality of the aesthetic of a news webpage was increased, the perception of bias in a news article was decreased. This effect was also common across all nine websites tested and across news articles with different levels of bias.

This contribution is important to the domain as it demonstrates that the visual presentation of a news website, through features and characteristics, can influence the perception of bias in a news article. Therefore, it is also possible, even probable, that other features and characteristics could be influencing other dimensions and measures of credibility such as expertise or trustworthiness. Future work should focus on identifying which other website design dimensions affect perceived bias or any other dimensions of credibility.

While individual features proved inconsistent influencers of perceived bias overall, it is likely that if the distortion effects were combined to make a more obvious or more severe visual effect that there would have been more consistency in the results. Nevertheless, the results do show that their inclusion or removal from a webpage can impact how bias is perceived in a news article. This experiment also identified suitable

characteristics of a news website's design or a webpage's aesthetic which should be investigated in the subsequent experiments.

8.2.2 Minor Contribution: The Experiment Framework and Platform, Including a Repository of Credibility Measures

The first minor contribution to the domain is the CAFE framework and accompanying experiment platform, including the repository of measures. This will make it easier for researchers to design, develop, and deploy experiments to measure credibility, or any of its individual dimensions and measures, on the WWW. It could also be used to aid in the design, development, and deployment of experiments measuring credibility or any of its dimensions and measures in other mediums. The CAFE framework was designed in response to the requirements identified in section 2.2.12.1 of this thesis.

The CAFE framework, which is detailed in section 3.3, shown in Figure 3-1, and its use described in section 3.3.2, is aimed at novice or new researchers to the domain. It is an open framework designed to be adapted to the researcher's individual experiment needs. The accompanying experiment platform, which is undergoing further development to add additional functionality, is aimed at making it easier for non-technical or novice researchers to deploy experiments. The platform was designed in response to the requirements identified in section 2.2.12.3.

One of the other major efforts of this thesis was the design and development of the repository and classification of measures used in peer reviewed academic research. This is further described in section 3.3.2. The table forms the first such repository and classification of its type in the domain. It enables researchers to quickly ascertain which measures have been used most often, how they have been used, and in what context. It enables them to search by the research application domain such as news, health or ecommerce, or by the dimension of credibility being measured such as source, medium or message. Researchers can quickly identify which measures are commonly used together, and other measures which may be useful to them. It also acts as a central repository and source of reference with links to the original research papers saving prospective researchers hours in search time. This repository is also undergoing redevelopment to expand the range of information it contains on each study and to introduce new functionality such as faceted search and data visualisations. This repository was designed in response to the requirements identified in section 2.2.12.2.

8.2.3 Minor Contribution: An In-depth Review of Existing Frameworks and the Underlying Theory in the Domain

The State of the Art chapter of this thesis contributes the second minor contribution to the collective body of knowledge. This includes the most comprehensive collection of definitions of the term in the literature, and the first classifications of bias that impact the production, dissemination, and consumption of news.

This section is now being prepared for publication as a journal article. This article may include the first model or models to explain how users arrive at and form judgements of bias.

8.3 *Future Work*

The completion of this thesis has provided several promising avenues for future work, many of which are being actively followed. These include:

- In the short term, a publication is being prepared on the benefits and use of the open CAFE framework. This will include promotion of an expanded and redeveloped experiment platform that was used to conduct the experiments contributing to this thesis and has been used by final year and Masters students.
- An experiment and paper are also being prepared to demonstrate the benefits of using Visual Analogue Scales (VAS) as a superior scale for measuring credibility online or any of its dimension and measures, such as bias. Currently most researchers typically employ Likert or Likert type scales. The issues with Likert scales and other scales are discussed in detail in sections 2.2.12 of this thesis. The benefits of VAS are discussed in detail in section 3.6.4. This paper will also include validity and reliability tests of the scales for credibility and some of its primary dimensions.
- A journal paper is also being prepared on the evolution of the ten human-centric Models, Theories, Frameworks, and Schematics of credibility judgements online, including their underlying theory, namely the ELM, HSM, and the C&AHIP. This will include the first realisation of the HSM model. This will primarily be made up of section 2.2.8 and section 0 of the State of the Art.
- Another area of future endeavour will be the completion and publication of the classification of biases that impact the production, dissemination, and consumption of news, for each medium. The lack of such a clear categorisation to define, delineate, diagnose and record the different biases is one of the main issues holding back research in the domain.
 - Biases which impact the production of news are presented in section 2.3.4 and continued in sections 9.3 of the appendix. This includes Table 9-2, Table 9-3, Table 9-4, and Table 9-5 which provide an overview the global, domain, organisational, and journalistic and editorial biases which impact the production of news.
 - Biases which impact the dissemination of news in traditional mediums (print, radio, and television) are highlighted in section 2.3.5 of the thesis and are continued in detail in section 9.4 of the appendix. This includes Table 9-6, Table 9-7, and Table 9-8 which provide an overview of the biases that have been studied in the dissemination of news in radio, print, and television.
 - Biases which impact the dissemination of news on the WWW are highlighted in section 2.3.6 of the thesis.
- To demonstrate the issues with defining bias, an extensive collection of existing definitions from the literature has been collected to show their lack of cohesion. To the best of the authors

knowledge, this is the most comprehensive collection existing in the literature. This has been classified by domain such as Definitions of Media Bias, or Definitions of Political Bias etc. These are presented in section 2.3.3 of this thesis and continued in detail in section 9.2 of the appendix.

- There is also significant scope to expand the online categorisation and repository of credibility measures. This is already under redevelopment to add additional functionality including a database back end, which will enable faceted search and the creation of visualisations. Publications will be prepared to promote its existence, benefits and use. They will also include analysis of how different dimensions and measures have been utilised over time in different domains.
- Long term, one of the primary aims is the development of a model, or models, to explain how consumers arrive at and make judgements of bias when consuming news on the WWW or in other mediums. There is also an opportunity to create other models to explain judgements of bias in other areas such as sports commentary, economic reporting, or commercial sales.

8.4 Final Remarks

Many of the most prominent researchers in the domain have lamented that bias has received surprisingly little empirical or theoretical study, despite all the furore it incites amid accusations and counter accusations, and its strategic use as an invective by many to describe information they simply disagree with. While there has been some attempt to remedy this situation with a range of empirical investigations into bias in the dissemination of print and television news, bias in the dissemination of news on the WWW continues to receive limited empirical attention. Armed with the knowledge that judgements of the credibility of information online is affected by its visual presentation, and with the knowledge that bias is a core dimension and measure of credibility, especially when forming judgements of bias in online news, this research was conducted under the supposition that judgements of bias in online news was similarly affected. It demonstrated that individual features, and in particular the wider characteristics of a news website's design, reflected in a news webpage's aesthetic, impacts the perception of bias, and what is more, that it can be predictively influenced. This demonstrates that judgements of bias, like judgements of credibility, are at least partially or initially judged heuristically. This will have a significant impact of the development of theoretical models to explain how users form judgements of bias online, a significant gap in the body of knowledge.

Bibliography

- Abdi, H. (2007). The Bonferroni and Šidák Corrections for Multiple Comparisons. In N. J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics. Vol. 1* (Vol. 1, pp. 103–107). SAGE Publications, Inc. <http://www.encyclopedias.biz/dw/Encyclopedia%20of%20Measurement%20and%20Statistics.pdf>
- Abdulla, R. A., Garrison, B., Salwen, M., Driscoll, P., & Casey, D. (2002). The credibility of newspapers, television news, and online news. *Artículo Presentado En La Association for Education in Journalism and Mass Communication*. <http://www.com.miami.edu/car/miamibeach1.pdf>
- Abraham, L., & Appiah, O. (2006). Framing News Stories: The Role of Visual Imagery in Priming Racial Stereotypes. *Howard Journal of Communications*, 17(3), 183–203. <https://doi.org/10.1080/10646170600829584>
- Aday, S. (2010). Chasing the Bad News: An Analysis of 2005 Iraq and Afghanistan War Coverage on NBC and Fox News Channel. *Journal of Communication*, 60(1), 144–164. <https://doi.org/10.1111/j.1460-2466.2009.01472.x>
- Addington, D. W. (1971). The effect of vocal variations on ratings of source credibility. *Speech Monographs*, 38(3), 242–247. <https://doi.org/10.1080/03637757109375716>
- Ahmed, A., Teo, C. H., Vishwanathan, S. V. N., & Smola, A. (2012). Fair and Balanced: Learning to Present News Stories. *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, 333–342. <https://doi.org/10.1145/2124295.2124337>
- Aitken, R. C. (1969). Measurement of feelings using visual analogue scales. *Proceedings of the Royal Society of Medicine*, 62(10), 989–993.
- Aitkin, D. (1972). Perceptions of partisan bias in the Australian mass media. *Politics*, 7(2), 160–169. <https://doi.org/10.1080/00323267208401290>
- Albert, W., & Tullis, T. (2013). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Newnes.
- Allcott, H., & Gentzkow, M. (2017). *Social Media and Fake News in the 2016 Election* (Working Paper No. 23089). National Bureau of Economic Research. <https://doi.org/10.3386/w23089>
- Allport, C. D., & Pendley, J. A. (2010). The impact of website design on the perceived credibility of internet financial reporting. *Intelligent Systems in Accounting, Finance & Management*, 17(3–4), 127–141. <https://doi.org/10.1002/isaf.318>
- Alsudani, F., & Casey, M. (2009). The Effect of Aesthetics on Web Credibility. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, 512–519. <http://dl.acm.org/citation.cfm?id=1671011.1671077>
- An, S., & Bergen, L. (2007). Advertiser Pressure on Daily Newspapers: A Survey of Advertising Sales Executives. *Journal of Advertising*, 36(2), 111–121. <https://doi.org/10.2753/JOA0091-3367360208>
- Ananny, M., & Crawford, K. (2014, September 10). Designer or journalist: Who shapes the news you read in your favorite apps? *Nieman Lab*. <https://www.niemanlab.org/2014/09/designer-or-journalist-who-shapes-the-news-you-read-in-your-favorite-apps/>
- Ananny, M., & Crawford, K. (2015). A Liminal Press. *Digital Journalism*, 3(2), 192–208. <https://doi.org/10.1080/21670811.2014.922322>
- Ansolabehere, S., Lessem, R., & Snyder, J. (2006). The Orientation of Newspaper Endorsements in U.S. Elections, 1940–2002. *Quarterly Journal of Political Science*, 393–404. <https://doi.org/10.1561/100.00000009>
- Appelman, A., & Sundar, S. S. (2016). Measuring Message Credibility: Construction and Validation of an Exclusive Scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>
- Arant, M. D., & Anderson, J. Q. (2001). Newspaper Online Editors Support Traditional Standards. *Newspaper Research Journal*, 22(4), 57–69. <https://doi.org/10.1177/073953290102200406>

- Aranyi, G., & Van Schaik, P. (2015). Modeling user experience with news websites. *Journal of the Association for Information Science and Technology*, 66(12), 2471–2493.
- Aristotle. (2015). *Rhetoric* (W. R. Roberts, Trans.). CreateSpace Independent Publishing Platform.
- Ayeh, J. K., Au, N., & Law, R. (2013). “Do We Believe in TripAdvisor?” Examining Credibility Perceptions and Online Travelers’ Attitude toward Using User-Generated Content. *Journal of Travel Research*, 0047287512475217. <https://doi.org/10.1177/0047287512475217>
- Ayton, P., & Tumber, H. (2001). The rise and fall of perceived bias at the BBC. *International Institute of Communications*, 29(4), 12–15.
- Babad, E. (1999). Preferential Treatment in Television Interviewing: Evidence from Nonverbal Behavior. *Political Communication*, 16(3), 337–358. <https://doi.org/10.1080/105846099198668>
- Baker, P. (2010). Representations of Islam in British broadsheet and tabloid newspapers 1999–2005. *Journal of Language and Politics*, 9(2), 310–338. <https://doi.org/10.1075/jlp.9.2.07bak>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132.
- Banning, S. A., & Sweetser, K. D. (2007). How Much Do They Think It Affects Them and Whom Do They Believe?: Comparing the Third-Person Effect and Credibility of Blogs and Traditional Media. *Communication Quarterly*, 55(4), 451–466. <https://doi.org/10.1080/01463370701665114>
- Banning, S., & Coleman, R. (2009). Louder than Words: A Content Analysis of Presidential Candidates’ Televised Nonverbal Communication. *Visual Communication Quarterly*, 16(1), 4–17. <https://doi.org/10.1080/15551390802620464>
- Bar-Ilan, J., Keenoy, K., Levene, M., & Yaari, E. (2009). Presentation bias is significant in determining user preference for search results—A user study. *Journal of the American Society for Information Science and Technology*, 60(1), 135–149. <https://doi.org/10.1002/asi.20941>
- Barker, D. C. (2002). *Rushed to Judgment: Talk Radio, Persuasion, and American Political Behavior*. Columbia University Press.
- Barnes, S., & Vidgen, R. (2000). WebQual: An Exploration of Website Quality. *Proceedings of the European Conference on Information Systems*, 9.
- Barnett, B. (2003). Guilty and Threatening: Visual Bias in Television News Crime Stories. *Journalism & Communication Monographs*, 5(3), 104–155. <https://doi.org/10.1177/152263790300500301>
- Baron, D. P. (2006). Persistent media bias. *Journal of Public Economics*, 90(1), 1–36. <https://doi.org/10.1016/j.jpubeco.2004.10.006>
- Barrett, A. W., & Barrington, L. W. (2005). Bias in newspaper photograph selection. *Political Research Quarterly*, 58(4), 609–618.
- Bartels, L. M. (2002). Beyond the Running Tally: Partisan Bias in Political Perceptions. *Political Behavior*, 24(2), 117–150. <https://doi.org/10.1023/A:1021226224601>
- Bassett, R. E., Staton-Spicer, A. Q., & Whitehead, J. L. (1979). Effects of source attire on judgments of credibility. *Central States Speech Journal*, 30(3), 282–285. <https://doi.org/10.1080/10510977909368022>
- Batlin, R. (1954). San Francisco Newspapers’ Campaign Coverage: 1896, 1952. *Journalism Bulletin*, 31(3), 297–303. <https://doi.org/10.1177/107769905403100302>
- Baum, M. A., & Gussin, P. (2008). In the Eye of the Beholder: How Information Shortcuts Shape Individual Perceptions of Bias in the Media. *Quarterly Journal of Political Science*, 3(1), 1–38. <https://doi.org/10.1561/100.00007010>
- Baum, M., & Just, M. (2009). Bandwagon and Underdog Effects in the 2008 Presidential Primary Campaign: A Survey Experiment. *APSA 2009 (American Political Science Association) Toronto Meeting Paper*, 22.

- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43(3), 800–813. <https://doi.org/10.3758/s13428-011-0081-0>
- Belanche, D., Casaló, L. V., Flavián, C., & Schepers, J. (2014). Trust transfer in the continued usage of public e-services. *Information & Management*, 51(6), 627–640. <https://doi.org/10.1016/j.im.2014.05.016>
- Bélanger, F., & Carter, L. (2008). Trust and risk in e-government adoption. *The Journal of Strategic Information Systems*, 17(2), 165–176. <https://doi.org/10.1016/j.jsis.2007.12.002>
- Belin, L. (2002). The Russian Media in the 1990s. *Journal of Communist Studies and Transition Politics*, 18(1), 139–160. <https://doi.org/10.1080/13523270209696371>
- Bente, G., Baptist, O., & Leuschner, H. (2012). To buy or not to buy: Influence of seller photos and reputation on buyer trust and purchase behavior. *International Journal of Human-Computer Studies*, 70(1), 1–13. <https://doi.org/10.1016/j.ijhcs.2011.08.005>
- Berghel, H. (2017). Oh, What a Tangled Web: Russian Hacking, Fake News, and the 2016 US Presidential Election. *Computer*, 50(9), 87–91. <https://doi.org/10.1109/MC.2017.3571054>
- Berlo, D. K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for Evaluating the Acceptability of Message Sources. *Public Opinion Quarterly*, 33(4), 563–576. <https://doi.org/10.1086/267745>
- Bernstein, D. (2004). *Israel in the Media: A Guide to Producing Effective Media Critiques*. The American Jewish Committee. <http://www.kintera.org/atf/cf/%7B42D75369-D582-4380-8395-D25925B85EAF%7D/Israel%20in%20the%20Media.pdf>
- Berry, C., Scheffler, A., & Goldstein, C. (1993). Effects of text structure on the impact of heard news. *Applied Cognitive Psychology*, 7(5), 381–395. <https://doi.org/10.1002/acp.2350070503>
- Bhattacharjee, A. (2002). Individual Trust in Online Firms: Scale Development and Initial Test. *Journal of Management Information Systems*, 19(1), 211–241. <https://doi.org/10.1080/07421222.2002.11045715>
- Billings, A. C., & Tambosi, F. (2004). Portraying the United States vs portraying a champion: US network bias in the 2002 World Cup. *International Review for the Sociology of Sport*, 39(2), 157–165.
- Bolsen, T., Druckman, J. N., & Cook, F. L. (2014). How Frames Can Undermine Support for Scientific Adaptations: Politicization and the Status-Quo Bias. *Public Opinion Quarterly*, 78(1), 1–26. <https://doi.org/10.1093/poq/nft044>
- Borah, P. (2014). The Hyperlinked World: A Look at How the Interactions of News Frames and Hyperlinks Influence News Credibility and Willingness to Seek Information. *Journal of Computer-Mediated Communication*, 19(3), 576–590. <https://doi.org/10.1111/jcc4.12060>
- Bovitz, G. L., Druckman, J. N., & Lupia, A. (2002). When Can a News Organization Lead Public Opinion? – Ideology Versus Market Forces in Decisions to Make News. *Public Choice*, 113(1–2), 127–155. <https://doi.org/10.1023/A:1020350716201>
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). *Is the Internet Causing Political Polarization? Evidence from Demographics* (Working Paper No. 23258). National Bureau of Economic Research. <https://doi.org/10.3386/w23258>
- Boyd, M. (2016). *Lead Stories and Elections: Prominence of News Coverage of Recent Election Campaigns in Four Liberal Media Markets*. 24.
- Bracken, C. C. (2006). Perceived Source Credibility of Local Television News: The Impact of Television Form and Presence. *Journal of Broadcasting & Electronic Media*, 50(4), 723–741. https://doi.org/10.1207/s15506878jobem5004_9
- Brandenburg, H. (2005). Political Bias in the Irish Media: A Quantitative Study of Campaign Coverage during the 2002 General Election. *Irish Political Studies*, 20(3), 297–322. <https://doi.org/10.1080/07907180500359350>
- Brandenburg, H. (2006). Party Strategy and Media Bias: A Quantitative Analysis of the 2005 UK Election Campaign. *Journal of Elections, Public Opinion and Parties*, 16(2), 157–178. <https://doi.org/10.1080/13689880600716027>

- Brandon, J. R. (1956). An experimental television study: The relative effectiveness of presenting factual information by the lecture, interview, and discussion methods. *Speech Monographs*, 23(4), 272–283. <https://doi.org/10.1080/03637755609375184>
- Brann, M., & Himes, K. L. (2010). Perceived Credibility of Male Versus Female Television Newscasters. *Communication Research Reports*, 27(3), 243–252. <https://doi.org/10.1080/08824091003737869>
- Brennan, E. (2016). Television in Ireland before Irish Television: Nationalist Rhetoric and International Programming. *Proceedings of ECREA*, 17.
- Brodie, M., Brady, L. A., & Altman, D. E. (1998). Media coverage of managed care: Is there a negative bias? *Health Affairs*, 17(1), 9–25. <https://doi.org/10.1377/hlthaff.17.1.9>
- Brohman, J. (1995). Universalism, Eurocentrism, and Ideological Bias in Development Studies: From Modernisation to Neoliberalism. *Third World Quarterly*, 16(1), 121–140.
- Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, 87(7), 1801–1824. [https://doi.org/10.1016/S0047-2727\(01\)00186-4](https://doi.org/10.1016/S0047-2727(01)00186-4)
- Bruter, M. (2009). Time Bomb?: The Dynamic Effect of News and Symbols on the Political Identity of European Citizens. *Comparative Political Studies*, 42(12), 1498–1536. <https://doi.org/10.1177/0010414009332465>
- Buckley, J. W., Buckley, M. H., & Chiang, H.-F. (1976). *Research Methodology and Business Decisions*.
- Bucy, E. P. (2003). Media Credibility Reconsidered: Synergy Effects between On-Air and Online News. *Journalism & Mass Communication Quarterly*, 80(2), 247–264. <https://doi.org/10.1177/107769900308000202>
- Budak, C., Goel, S., & Rao, J. M. (2016). Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis. *Public Opinion Quarterly*, 80(S1), 250–271. <https://doi.org/10.1093/poq/nfw007>
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal Deception Theory. *Communication Theory*, 6(3), 203–242. <https://doi.org/10.1111/j.1468-2885.1996.tb00127.x>
- Bureau, U. C. (2016, July 16). *Congressional Voting Turnout is at Lowest Mark Since 1978*. <https://www.census.gov/newsroom/press-releases/2015/cb15-122.html>
- Burgess, A. (1997). Writing off Slovakia to “the east”? Examining charges of bias in British press reporting of Slovakia, 1993–1994. *Nationalities Papers*, 25(4), 659–682. <https://doi.org/10.1080/00905999708408533>
- Burgoon, J. K. (1978). Attributes of the Newscaster’s Voice as Predictors of His Credibility. *Journalism Quarterly*, 55(2), 276–300. <https://doi.org/10.1177/107769907805500208>
- Burgoon, J. K., & Saine, T. P. (1978). *The unspoken dialogue: An introduction to nonverbal communication*. Houghton Mifflin Harcourt (HMH).
- Burke, M. A., & Eichler, M. (2006). *The BIAS FREE Framework. A practical tool for identifying and eliminating social biases in health research*. Global Forum for Health Research. <https://tspace.library.utoronto.ca/handle/1807/9581>
- Burkell, J. (2004). Health Information Seals of Approval: What do they Signify? *Information, Communication & Society*, 7(4), 491–509. <https://doi.org/10.1080/1369118042000305610>
- Cacioppo, J. T., & Petty, R. E. (1989). Effects of message repetition on argument processing, recall, and persuasion. *Basic and Applied Social Psychology*, 10(1), 3–12.
- Callan, N. (2017, September 6). *New Strategic Communications Unit Established* [Government News Service]. MerrionStreet.Ie - Irish Government News Service. https://merrionstreet.ie/en/News-Room/News/New_Strategic_Communications_Unit_Established.html
- Carter, R. F., & Greenberg, B. S. (1965). Newspapers or Television: Which Do You Believe? *Journalism Quarterly*, 42(1), 29–34. <https://doi.org/10.1177/107769906504200104>
- Carter, T., Franklin, M., Sanders, A., & Wright, J. (2011). *The First Amendment and the Fourth Estate: The Law of Mass Media* (11 edition). Foundation Press.

- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's mTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- Cassidy, W. P. (2006). Gatekeeping similar for online, print journalists. *Newspaper Research Journal*, 27(2), 6.
- Cassidy, W. P. (2007). Online News Credibility: An Examination of the Perceptions of Newspaper Journalists. *Journal of Computer-Mediated Communication*, 12(2), 478–498. <https://doi.org/10.1111/j.1083-6101.2007.00334.x>
- Castillo, C., Mendoza, M., & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. *Internet Research*, 23(5), 560–588. <https://doi.org/10.1108/IntR-05-2012-0095>
- Castillo, C., Mendoza, M., & Poblete, B. (2011). Information Credibility on Twitter. *Proceedings of the 20th International Conference on World Wide Web*, 675–684. <https://doi.org/10.1145/1963405.1963500>
- Ceplair, L. (2008). Reporting The Blacklist: Anti-Communist Challenges To Elizabeth Poe Kerby. *Historical Journal of Film, Radio and Television*, 28(2), 135–151. <https://doi.org/10.1080/01439680802077139>
- Chaiken, S. (1977). *The Use of Source Versus Message Cues in Persuasion: An Information Processing Analysis* [Doctor of Philosophy, University of Massachusetts - Amherst]. https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=2657&context=dissertations_1
- Chaiken, S. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766.
- Chaiken, S., & Eagly, A. H. (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology*, 34(4), 605–614.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process Theories in Social Psychology*. Guilford Press.
- Chakraborty, A., Ghosh, S., Ganguly, N., & Gummadi, K. P. (2015). Can Trending News Stories Create Coverage Bias? On the Impact of High Content Churn in Online News Media. *Computation and Journalism Symposium*, 5.
- Chakraborty, A., Ghosh, S., Ganguly, N., & Gummadi, K. P. (2016). Dissemination Biases of Social Media Channels: On the Topical Coverage of Socially Shared News. *ICWSM*, 559–562.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, 46(1), 112–130. <https://doi.org/10.3758/s13428-013-0365-7>
- Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. *Information & Management*, 42(4), 543–559. <https://doi.org/10.1016/j.im.2004.02.006>
- Cheung, C., & Lee, M. (2000). Trust in Internet Shopping: A Proposed Model and Measurement. *Instrument," Proceedings of the 2000 AMCIS Conference*.
- Cheung, C., & Lee, M. (2001). Trust in Internet Shopping: Instrument Development and Validation through Classical and Modern Approaches. *Journal of Global Information Management (JGIM)*, 9(3), 23–35.
- Chiang, C.-F., & Knight, B. (2011). Media Bias and Influence: Evidence from Newspaper Endorsements. *The Review of Economic Studies*, 78(3), 795–820. <https://doi.org/10.1093/restud/rdq037>
- Choi, S. M., & Rifon, N. J. (2002). Antecedents and Consequences of Web Advertising Credibility. *Journal of Interactive Advertising*, 3(1), 12–24. <https://doi.org/10.1080/15252019.2002.10722064>
- Choi, W., & Stvilia, B. (2015a). *A New Framework for Web Credibility Assessment*. <https://www.ideals.illinois.edu/handle/2142/73738>
- Choi, W., & Stvilia, B. (2015b). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, 66(12), 2399–2414. <https://doi.org/10.1002/asi.23543>
- Chu, S.-C., & Kamal, S. (2008). The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes: An exploratory study. *Journal of Interactive Advertising*, 8(2), 26–37.

- Chung, C. J., Kim, H., & Kim, J. H. (2010). An anatomy of the credibility of online newspapers. *Online Information Review*, 34(5), 669–685. <https://doi.org/10.1108/14684521011084564>
- Chung, C. J., Nam, Y., & Stefanone, M. A. (2012). Exploring Online News Credibility: The Relative Influence of Traditional and Technological Factors. *Journal of Computer-Mediated Communication*, 17(2), 171–186. <https://doi.org/10.1111/j.1083-6101.2011.01565.x>
- Cockburn, A., Gutwin, C., & Dix, A. (2018). HARK No More: On the Preregistration of CHI Experiments. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 141:1–141:12. <https://doi.org/10.1145/3173574.3173715>
- Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. *Proceedings of the British Educational Research Association Annual Conference*, 18. <http://www.leeds.ac.uk/educol/documents/00002182.htm>
- Cohen, B. C. (1963). *The Press and Foreign Policy*. Institute of Governmental Studies, University of California.
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. Academic Press.
- Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58(6), 737–758. [https://doi.org/10.1016/S1071-5819\(03\)00041-7](https://doi.org/10.1016/S1071-5819(03)00041-7)
- Corritore, C. L., Marble, R. P., Wiedenbeck, S., Kracher, B., & Chandran, A. (2005). Measuring online trust of websites: Credibility, perceived ease of use, and risk. *AMCIS 2005 Proceedings*, 370.
- Cox, D. R., & Reid, N. (2000). *The theory of the design of experiments*. CRC Press.
- Coyne, E. (2018a, February 23). Government 'paid for good news stories' over Ireland 2040 plan. *The Times*. <https://www.thetimes.co.uk/article/martin-attacks-dubious-2040-media-deal-c6wtq772d>
- Coyne, E. (2018b, February 26). Make 2040 ads look like real news, papers told | Ireland | The Times. *The Times*. <https://www.thetimes.co.uk/article/make-2040-ads-look-like-real-news-papers-told-kq20xcs67>
- Coyne, E. (2018c, February 28). Ireland 2040 newspapers told there would be 'more to come.' *The Times*. <https://www.thetimes.co.uk/article/ireland-2040-newspapers-told-there-would-be-more-to-come-sc6gjp88g>
- Crandall, C. S., Eidelman, S., Skitka, L. J., & Morgan, G. S. (2009). Status quo framing increases support for torture. *Social Influence*, 4(1), 1–10. <https://doi.org/10.1080/15534510802124397>
- Craswell, N., Zoeter, O., Taylor, M., & Ramsey, B. (2008). An Experimental Comparison of Click Position-bias Models. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 87–94. <https://doi.org/10.1145/1341531.1341545>
- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Cugelman, B., Thelwall, M., & Dawes, P. L. (2009). Dimensions of web site credibility and their relation to active trust and behavioural impact. *Communications of the Association for Information Systems*, 24, 455–472.
- Cuthbert, M. (1976). Some Observations on the Role of Mass Media in the Recent Socio-Political Development of Jamaica. *Caribbean Quarterly*, 22(4), 50–58. JSTOR.
- Cyr, D. (2008). Modeling Web Site Design Across Cultures: Relationships to Trust, Satisfaction, and E-Loyalty. *Journal of Management Information Systems*, 24(4), 47–72. <https://doi.org/10.2753/MIS0742-1222240402>
- Cyr, D., Head, M., & Larios, H. (2010). Colour appeal in website design within and across cultures: A multi-method evaluation. *International Journal of Human-Computer Studies*, 68(1–2), 1–21. <https://doi.org/10.1016/j.ijhcs.2009.08.005>
- D'Alessio, D., & Allen, M. (2000). Media bias in presidential elections: A meta-analysis. *Journal of Communication*, 50(4), 133–156. <https://doi.org/10.1111/j.1460-2466.2000.tb02866.x>

- D'Alessio, D. W., & Allen, M. (2007). On the role of newspaper ownership on bias in presidential campaign coverage by newspapers. In R. W. Preiss, B. M. Gayle, N. Burrell, & J. Bryant (Eds.), *Mass Media Effects Research: Advances Through Meta-analysis* (p. 429). Psychology Press.
- Dalton, R. J., Beck, P. A., & Huckfeldt, R. (1998). Partisan Cues and the Media: Information Flows in the 1992 Presidential Election. *American Political Science Review*, 111–126.
- D'Angelo, P., & Kuypers, J. A. (2010). *Doing News Framing Analysis: Empirical and Theoretical Perspectives*. Routledge.
- Davis, C. J., Bowers, J. S., & Memon, A. (2011). Social Influence in Televised Election Debates: A Potential Distortion of Democracy. *PLoS One (Public Library of Science Journal)*, 6(3), e18154. <https://doi.org/10.1371/journal.pone.0018154>
- Deacon, D., Golding, P., & Billig, M. (2001). Press and Broadcasting: 'Real Issues' and Real Coverage. *Parliamentary Affairs*, 54(4), 666–678. <https://doi.org/10.1093/parlij/54.4.666>
- Debenport, E. (1992, September 27). Now about that media bias. *St Petersburg Times*, 1D.
- DellaVigna, S., & Kaplan, E. (2006). *The Fox News effect: Media bias and voting*. National Bureau of Economic Research.
- DellaVigna, Stefano, & Kaplan, E. (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3), 1187–1234.
- DellaVigna, Stefano, & Kennedy, A. (2011). *Does Media Concentration Lead to Biased Coverage? Evidence from Movie Reviews**.
- Dennis, E. E. (1997). How “liberal” are the media, anyway? The continuing conflict of professionalism and partisanship. *Harvard International Journal of Press/Politics*, 2(4), 115–119.
- Desmond, R., & Danilewicz, A. (2010). Women Are On, But Not In, the News: Gender Roles in Local Television News. *Sex Roles*, 62(11–12), 822–829. <https://doi.org/10.1007/s11199-009-9686-5>
- Desmond, R. W. (1949). Book Review: Royal Commission on the Press, 1947–1949: Report. *Journalism Bulletin*, 26(4), 470–472. <https://doi.org/10.1177/107769904902600423>
- Di Massa, R., Montagnuolo, M., & Messina, A. (2010). Implicit News Recommendation Based on User Interest Models and Multimodal Content Analysis. *Proceedings of the 3rd International Workshop on Automated Information Extraction in Media Production*, 33–38. <https://doi.org/10.1145/1877850.1877861>
- Di Tella, R., & Franceschelli, I. (2011). Government Advertising and Media Coverage of Corruption Scandals. *American Economic Journal: Applied Economics*, 3(4), 119–151.
- Dick, M. (2015). Just Fancy That: An analysis of infographic propaganda in The Daily Express, 1956–1959. *Journalism Studies*, 16(2), 152–174. <https://doi.org/10.1080/1461670X.2013.872415>
- Diddi, A., Fico, F., & Zeldes, G. A. (2014). Partisan Balance and Bias in TV Network Coverage of the 2000, 2004, and 2008 Presidential Elections. *Journal of Broadcasting & Electronic Media*, 58(2), 161–178. <https://doi.org/10.1080/08838151.2014.906431>
- Dingli, A., & Cassar, S. (2014). An intelligent framework for website usability. *Advances in Human-Computer Interaction*, 2014.
- Dingli, A., & Mifsud, J. (2011). *Useful: A framework to mainstream web site usability through automated evaluation*.
- Doll, H. D., & Bradley, B. E. (1974). A study of the objectivity of television news reporting of the 1972 presidential campaign. *Central States Speech Journal*, 25(4), 254–263. <https://doi.org/10.1080/10510977409367803>
- Dominick, J. R. (1977). Geographic Bias in National TV News. *Journal of Communication*, 27(4), 94–99. <https://doi.org/10.1111/j.1460-2466.1977.tb01862.x>
- Donsbach, W. (2004). Psychology of News Decisions: Factors behind Journalists' Professional Behavior. *Journalism*, 5(2), 131–157. <https://doi.org/10.1177/146488490452002>
- Donsbach, W., & Klett, B. (1993). Subjective objectivity. How journalists in four countries define a key term of their profession. *Gazette (Leiden, Netherlands)*, 51(1), 53–83. <https://doi.org/10.1177/001654929305100104>

- Druckman, J. N., & Parkin, M. (2005). The Impact of Media Bias: How Editorial Slant Affects Voters. *The Journal of Politics*, 67(4), 1030–1049.
- DuBois, P. (2008). *MySQL*. Pearson Education.
- Earl, J., Martin, A., McCarthy, J. D., & Soule, S. A. (2004). The Use of Newspaper Data in the Study of Collective Action. *Annual Review of Sociology*, 30, 65–80.
- Eastin, M. S. (2006). Credibility Assessments of Online Health Information: The Effects of Source Expertise and Knowledge of Content. *Journal of Computer-Mediated Communication*, 6(4), 0–0. <https://doi.org/10.1111/j.1083-6101.2001.tb00126.x>
- Eastman, S. T., & Billings, A. C. (2000). Sportscasting and Sports Reporting: The Power of Gender Bias. *Journal of Sport and Social Issues*, 24(2), 192–213. <https://doi.org/10.1177/0193723500242006>
- Efron, E. (1971). *The News Twisters*. Nash Pub.
- Egger, F. N. (2001). Affective Design of E-Commerce User Interfaces. *Proceedings of The International Conference on Affective Human Factors Design*, 8.
- Egger, F. N. (2000). “Trust Me, I’m an Online Vendor”: Towards a Model of Trust for E-Commerce System Design. *CHI '00 Extended Abstracts on Human Factors in Computing Systems.*, 2.
- Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering Into the Bias Blind Spot: People’s Assessments of Bias in Themselves and Others. *Personality and Social Psychology Bulletin*, 31(5), 680–692. <https://doi.org/10.1177/0146167204271570>
- Eisend, M. (2006). Source Credibility Dimensions in Marketing Communication—A Generalized Solution. *Journal of Empirical Generalisations in Marketing*, 10(2), 33.
- Eisinger, R. M., Veenstra, L. R., & Koehn, J. P. (2007). What Media Bias? Conservative and Liberal Labeling in Major U.S. Newspapers. *Harvard International Journal of Press/Politics*, 12(1), 17–36. <https://doi.org/10.1177/1081180X06297460>
- Ekström, M., Eriksson, G., Johansson, B., & Wikström, P. (2013). Biased Interrogations? *Journalism Studies*, 14(3), 423–439. <https://doi.org/10.1080/1461670X.2012.689488>
- Ellis, G. (2014). *Trust Ownership and the Future of News: Media Moguls and White Knights*. Palgrave Macmillan UK. <https://www.palgrave.com/gp/book/9781137369437>
- Emerson, J. W., Seltzer, M., & Lin, D. (2009). Assessing judging bias: An example from the 2000 Olympic Games. *The American Statistician*, 63(2), 124–131.
- Entman, R. M. (1990). Modern racism and the images of blacks in local television news. *Critical Studies in Mass Communication*, 7(4), 332–345. <https://doi.org/10.1080/15295039009360183>
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Entman, R. M. (2007). Framing Bias: Media in the Distribution of Power. *Journal of Communication*, 57(1), 163–173. <https://doi.org/10.1111/j.1460-2466.2006.00336.x>
- Epstein, E. J. (1973). *News from nowhere: Television and the news*. Random House.
- Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, 55, 62–73. <https://doi.org/10.1016/j.tourman.2016.01.013>
- Esser, F. (2008). Dimensions of Political News Cultures: Sound Bite and Image Bite News in France, Germany, Great Britain, and the United States. *The International Journal of Press/Politics*, 13(4), 401–428. <https://doi.org/10.1177/1940161208323691>
- Esser, F., Stepińska, A., & Hopmann, D. N. (2016). Populism and the Media. Cross-National Findings and Perspectives. In T. Aalberg, F. Esser, C. Reinemann, J. Strömbäck, & C. H. de Vreese (Eds.), *Populist Political Communication in Europe* (pp. 365–381). Routledge. <https://doi.org/10.4324/9781315623016>

- Evans, J. S. B. (1989). *Bias in Human Reasoning Causes and Consequences*.
- Eveland, W. P., & Shah, D. V. (2003). The Impact of Individual and Interpersonal Factors on Perceived News Media Bias. *Political Psychology, 24*(1), 101–117. <https://doi.org/10.1111/0162-895X.00318>
- Ewing, T. N. (1942). A Study of Certain Factors Involved in Changes of Opinion. *The Journal of Social Psychology, 16*(1), 63–88. <https://doi.org/10.1080/00224545.1942.9714105>
- Eysenbach, G. (2008). Credibility of health information and digital media: New perspectives and implications for youth. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 123–154). MIT Press. <https://tspace.library.utoronto.ca/handle/1807/9906>
- Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ, 324*(7337), 573–577. <https://doi.org/10.1136/bmj.324.7337.573>
- Fahmy, S., & Johnson, T. J. (2005). “How we Performed”: Embedded Journalists’ Attitudes and Perceptions Towards Covering the Iraq War. *Journalism & Mass Communication Quarterly, 82*(2), 301–317. <https://doi.org/10.1177/107769900508200205>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Feldman, L. (2011). The Effects of Journalist Opinionation on Learning From the News. *Journal of Communication, 61*(6), 1183–1201. <https://doi.org/10.1111/j.1460-2466.2011.01602.x>
- Fichter, C., & Jonas, K. (2008). Image Effects of Newspapers. *Zeitschrift Für Psychologie / Journal of Psychology, 216*(4), 226–234. <https://doi.org/10.1027/0044-3409.216.4.226>
- Fico, F., Richardson, J. D., & Edwards, S. M. (2004). Influence of Story Structure on Perceived Story Bias and News Organization Credibility. *Mass Communication and Society, 7*(3), 301–318. https://doi.org/10.1207/s15327825mcs0703_3
- Filak, V. F. (2004). Cultural Convergence: Intergroup Bias Among Journalists and its Impact on Convergence. *Atlantic Journal of Communication, 12*(4), 216–232. https://doi.org/10.1207/s15456889ajc1204_3
- File, T. (2015). Who Votes? Congressional Elections and the American Electorate: 1978–2014. *Census Bureau*. <http://www.census.gov/edgekey-staging.net/content/dam/Census/library/publications/2015/demo/p20-577.pdf>
- Finn, C. (2018, March 6). Leo admits “mistakes were made” and “controls were too loose” with his communications unit. *TheJournal.Ie*. <http://www.thejournal.ie/scu-leo-varadkar-leaders-questions-3888359-Mar2018/>
- Finn, S. M. (1982). *Mass Media and the 1981 Election: An Analysis of the Extent and Bias of Mass Media Coverage of the 1981 South African Election*. Tafelberg.
- Fisher, S. R. A. (1935). *The Design of Experiments*. Oliver and Boyd.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism & Mass Communication Quarterly, 77*(3), 515–540.
- Flanagin, A. J., & Metzger, M. J. (2003). The perceived credibility of personal Web page information as influenced by the sex of the source. *Computers in Human Behavior, 19*(6), 683–701. [https://doi.org/10.1016/S0747-5632\(03\)00021-9](https://doi.org/10.1016/S0747-5632(03)00021-9)
- Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society, 9*(2), 319–342. <https://doi.org/10.1177/1461444807075015>

- Fogg, B. J. (2003). Prominence-Interpretation Theory: Explaining how People Assess Credibility Online. *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 722–723. <https://doi.org/10.1145/765891.765951>
- Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2001). What makes Web sites credible?: A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 61–68. <https://doi.org/10.1145/365024.365037>
- Fogg, B. J., Marshall, J., Osipovich, A., Varma, C., Laraki, O., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2000). Elements That Affect Web Credibility: Early Results from a Self-report Study. *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, 287–288. <https://doi.org/10.1145/633292.633460>
- Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do Users Evaluate the Credibility of Web sites?: A Study with Over 2,500 Participants. *Proceedings of the 2003 Conference on Designing for User Experiences*, 1–15. <http://dl.acm.org/citation.cfm?id=997097>
- Fogg, B. J., & Tseng, H. (1999). The Elements of Computer Credibility. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 80–87. <https://doi.org/10.1145/302979.303001>
- Folstein, M. F., & Luria, R. (1973). Reliability, validity, and clinical application of the Visual Analogue Mood Scale. *Psychological Medicine*, 3(04), 479–486.
- Forward, R. (1977). Editorial opinion and the Whitlam government. *Politics*, 12(1), 136–141. <https://doi.org/10.1080/00323267708401596>
- Frank, R. S. (1973). *Message Dimensions of Television News*. Lexington Books.
- Freeman, H. E., Weeks, H. A., & Wertheimer, W. J. (1955). News Commentator Effect: A Study in Knowledge and Opinion Change. *Public Opinion Quarterly*, 19(2), 209–215.
- Freeman, K. S., & Spyridakis, J. H. (2004). An Examination of Factors That Affect the Credibility of Online Health Information. *Technical Communication*, 51(2), 239–263.
- Freeman, K. S., & Spyridakis, J. H. (2009). Effect of Contact Information on the Credibility of Online Health Information. *IEEE Transactions on Professional Communication*, 52(2), 152–166. <https://doi.org/10.1109/TPC.2009.2017992>
- Frewer, L. J., Howard, C., Hedderley, D., & Shepherd, R. (2006). What Determines Trust in Information About Food-Related Risks? Underlying Psychological Constructs. *Risk Analysis*, 16(4), 473–486. <https://doi.org/10.1111/j.1539-6924.1996.tb01094.x>
- Frewer, Lynn J., Howard, C., Hedderley, D., & Shepherd, R. (1997). The Elaboration Likelihood Model and Communication About Food Risks. *Risk Analysis*, 17(6), 759–770. <https://doi.org/10.1111/j.1539-6924.1997.tb01281.x>
- Friedman, H. S., Mertz, T. I., & DiMatteo, M. R. (1980). Perceived Bias in the Facial Expressions of Television News Broadcasters. *Journal of Communication*, 30(4), 103–111. <https://doi.org/10.1111/j.1460-2466.1980.tb02022.x>
- Friedman, S. (2011). Whose freedom? South Africa's press, middle-class bias and the threat of control. *Ecquid Novi: African Journalism Studies*, 32(2), 106–121. <https://doi.org/10.1080/02560054.2011.578887>
- Fritch, J. W., & Cromwell, R. L. (2001). Evaluating Internet resources: Identity, affiliation, and cognitive authority in a networked world. *Journal of The American Society for Information Science and Technology*, 52(6), 499–507. <https://doi.org/10.1002/asi.1081>
- Frost, C. (2015). *Journalism Ethics and Regulation* (4 edition). Routledge.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. Free Press Paperbacks.
- Gaber, I., Seymour, E., & Thomas, L. (2009). Review Commentary: Is the BBC biased? The Corporation and the coverage of the 2006 Israeli—Hezbollah war. *Journalism*, 10(2), 239–259.
- Gaddy, G. D., & Tanjong, E. (1986). Earthquake Coverage by the Western Press. *Journal of Communication*, 36(2), 105–112. <https://doi.org/10.1111/j.1460-2466.1986.tb01428.x>
- Gambaro, M., & Puglisi, R. (2015). What do ads buy? Daily coverage of listed companies on the Italian press. *European Journal of Political Economy*, 39, 41–57. <https://doi.org/10.1016/j.ejpoleco.2015.03.008>

- Gambetta, D. (2000). Can We Trust Trust? In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 213–237). Can we trust trust.
- Ganguly, B., Dash, S. B., Cyr, D., & Head, M. (2010). The effects of website design on purchase intention in online shopping: The mediating role of trust and the moderating role of culture. *International Journal of Electronic Business*, 8(4–5), 302–330. <https://doi.org/10.1504/IJEB.2010.035289>
- Gantz, W. (1981). The influence of researcher methods on television and newspaper news credibility evaluations. *Journal of Broadcasting*, 25(2), 155–169. <https://doi.org/10.1080/08838158109386439>
- Ganzaroli, A., Tan, Y.-H., & Thoen, W. (1999). The social and institutional context of trust in electronic commerce. *Autonomous Agents*, 99, 65–76.
- Gardner, J. (Ed.). (2009). The Web Server Gateway Interface (WSGI). In *The Definitive Guide to Pylons* (pp. 369–388). Apress. https://doi.org/10.1007/978-1-4302-0534-0_16
- Garforth, C., Grimble, R., Sutherland, A., Great Britain, Department for International Development, & Natural Resources Institute (Great Britain). (1998). *Socio-economic methodologies: Best practice guidelines*. Natural Resources Institute.
- Garrett, R. K., & Stroud, N. J. (2014). Partisan paths to exposure diversity: Differences in pro-and counterattitudinal news consumption. *Journal of Communication*, 64(4), 680–701.
- Gass, R. H., & Seiter, J. S. (2015). *Persuasion: Social Influence and Compliance Gaining*. Routledge.
- Gaziano, C., & McGrath, K. (1986). Measuring the concept of credibility. *Journalism Quarterly*, 63(3), 451–62. <http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ346970>
- Gaziano, Cecilie. (1988). How credible is the credibility crisis? *Journalism Quarterly*, 65, 267.
- Gaziano, Cecilie, & McGrath, K. (1986). Measuring the Concept of Credibility. *Journalism Quarterly*, 63(3). <http://search.proquest.com/docview/216931467/13A4707980A4C273036/1?accountid=14404>
- Gefen, D., & Straub, D. W. (2003). Managing User Trust in B2C e-Services. *E-Service Journal*, 2(2), 7–24.
- Gentzkow, M., & Shapiro, J. M. (2006). Media Bias and Reputation. *Journal of Political Economy*, 114, 280–316.
- Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1), 35–71. <https://doi.org/10.3982/ECTA7195>
- Gentzkow, M., Shapiro, J. M., & Stone, D. F. (2015). Chapter 14 - Media Bias in the Marketplace: Theory. In S. P. Anderson, J. Waldfogel, & D. Strömberg (Eds.), *Handbook of Media Economics* (Vol. 1, pp. 623–645). North-Holland. <https://doi.org/10.1016/B978-0-444-63685-0.00014-0>
- Gerber, A. S., Karlan, D., & Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, 1(2), 35–52. <https://doi.org/10.1257/app.1.2.35>
- Giffin, K. (1967). The Contribution Of Studies Of Source Credibility To A Theory Of Interpersonal Trust In The Communication Process. *Psychological Bulletin*, 68(2), 104–120.
- Gilens, M., & Hertzman, C. (2000). Corporate Ownership and News Bias: Newspaper Coverage of the 1996 Telecommunications Act. *The Journal of Politics*, 62(2), 369–386. <https://doi.org/10.1111/0022-3816.00017>
- Gilliam, F. D., Valentino, N. A., & Beckmann, M. N. (2002). Where You Live and What You Watch: The Impact of Racial Proximity and Local Television News on Attitudes about Race and Crime. *Political Research Quarterly*, 55(4), 755–780. <https://doi.org/10.1177/106591290205500402>
- Giudice, K. D. (2010). Crowdsourcing credibility: The impact of audience feedback on Web page credibility. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–9. <https://doi.org/10.1002/meet.14504701099>

- Gladney, G. A. (1996). How Editors and Readers Rank and Rate the Importance of Eighteen Traditional Standards of Newspaper Excellence. *Journalism & Mass Communication Quarterly*, 73(2), 319–331. <https://doi.org/10.1177/107769909607300204>
- GNM press office. (2016, July 21). Guardian reports outstanding digital and print ABC figures in June. *The Guardian*. <https://www.theguardian.com/gnm-press-office/2016/jul/21/guardian-reports-outstanding-digital-and-print-abc-figures-in-june>
- Goldberg, B. (2003). *Bias: A CBS Insider Exposes How the Media Distort the News*. HarperCollins.
- González-Bailón, S., Morales, G. D. F., Mendoza, M., Khan, N., & Castillo, C. (2014). Cable news coverage and online news stories: A large-scale comparison of digital media content. *Annual Meeting of the International Communication Association (ICA)*.
- Goodnow, T. (2010). Visual Bias in Time’s “The Great Divide”: A Semiotic Analysis of Clinton and Obama Photographs. *American Behavioral Scientist*, 54(4), 406–416. <https://doi.org/10.1177/0002764210381865>
- Gorham, B. W. (2006). News Media’s Relationship with Stereotyping: The Linguistic Intergroup Bias in Response to Crime News. *Journal of Communication*, 56(2), 289–308. <https://doi.org/10.1111/j.1460-2466.2006.00020.x>
- Gottfried, J., Barthel, M., & Mitchell, A. (2017). *Pew Research Center, Trump, Clinton Voters Divided in Their Main Source for Election News*. Pew Research Center. <http://www.journalism.org/2017/01/18/trump-clinton-voters-divided-in-their-main-source-for-election-news/>
- Grabe, M. E. (1996). The South African Broadcasting Corporation’s Coverage of the 1987 and 1989 Elections: The Matter of Visual Bias. *Journal of Broadcasting & Electronic Media*, 40, 153–179.
- Grabe, M. E., & Bucy, E. P. (2009). *Image Bite Politics: News and the Visual Framing of Elections*. Oxford University Press, USA.
- Grabe, M. E., Trager, K. D., Lear, M., & Rauch, J. (2006). Gender in Crime News: A Case Study Test of the Chivalry Hypothesis. *Mass Communication and Society*, 9(2), 137–163. https://doi.org/10.1207/s15327825mcs0902_2
- Grandison, T., & Sloman, M. (2000). A survey of trust in internet applications. *IEEE Communications Surveys Tutorials*, 3(4), 2–16. <https://doi.org/10.1109/COMST.2000.5340804>
- Green-Pedersen, C., Mortensen, P. B., & Thesen, G. (2017). The Incumbency Bonus Revisited: Causes and Consequences of Media Dominance. *British Journal of Political Science*, 47(1), 131–148. <https://doi.org/10.1017/S0007123415000022>
- Greenwood, K. (2005). Picturing presidents: A content analysis of photographs of Presidents from the Pictures of the Year. *International Communication Association*, 1–34.
- Greer, J. D. (2003). Evaluating the Credibility of Online Information: A Test of Source and Advertising Influence. *Mass Communication and Society*, 6(1), 11–28. https://doi.org/10.1207/S15327825MCS0601_3
- Groeling, T. (2008). Who’s the Fairest of them All? An Empirical Test for Partisan Bias on ABC, CBS, NBC, and Fox News. *Presidential Studies Quarterly*, 631–657.
- Groeling, T. (2013a). Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News. *Annual Review of Political Science*, 16(1), 129–151. <https://doi.org/10.1146/annurev-polisci-040811-115123>
- Groeling, T. (2013b). Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News. *Annual Review of Political Science*, 16(1), 129–151. <https://doi.org/10.1146/annurev-polisci-040811-115123>
- Groeling, T., & Kernell, S. (1998). Is Network News Coverage of the President Biased? *The Journal of Politics*, 60(4), 1063–1087. <https://doi.org/10.2307/2647731>
- Groseclose, T. (2011). *Left Turn: How Liberal Media Bias Distorts the American Mind*. St. Martin’s Press.
- Groseclose, T., & Milyo, J. (2005a). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237. <http://www.jstor.org/stable/25098770>

- Groseclose, T., & Milyo, J. (2005b). A social-science perspective on media bias. *Critical Review*, 17(3–4), 305–314. <https://doi.org/10.1080/08913810508443641>
- Groseclose, T., & Milyo, J. (2005c). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237. <https://doi.org/10.1162/003355305775097542>
- Groseclose, T., & Milyo, J. (2005d). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237. <http://www.jstor.org/stable/25098770>
- Gunter, B. (1979). Recall of Brief Television News Items: Effects of Presentation Mode, Picture Content and Serial Position. *Journal of Educational Television*, 5(2), 57–61. <https://doi.org/10.1080/0260741790050207>
- Gunter, B. (1980). Remembering Television News: Effects of Picture Content. *The Journal of General Psychology*, 102(1), 127–133. <https://doi.org/10.1080/00221309.1980.9920970>
- Gunter, B. (1997). *Measuring Bias on Television*. University of Luton Press.
- Gunther, A. C. (1992). Biased Press or Biased Public? Attitudes Toward Media Coverage of Social Groups. *Public Opinion Quarterly*, 56(2), 147–167.
- Gupta, A., Kumaraguru, P., Castillo, C., & Meier, P. (2014). TweetCred: Real-Time Credibility Assessment of Content on Twitter. *Social Informatics*, 228–243. https://doi.org/10.1007/978-3-319-13734-6_16
- Guzman, A. L. (2016). Evolution of News Frames During the 2011 Egyptian Revolution: Critical Discourse Analysis of Fox News's and CNN's Framing of Protesters, Mubarak, and the Muslim Brotherhood. *Journalism & Mass Communication Quarterly*, 93(1), 80–98. <https://doi.org/10.1177/1077699015606677>
- Hackett, R. A. (1984). Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Mass Communication*, 1(3), 229–259. <https://doi.org/10.1080/15295038409360036>
- Hackett, R. A. (1985). A Hierarchy of Access: Aspects of Source Bias in Canadian TV News. *Journalism Quarterly*, 62(2), 256–277. <https://doi.org/10.1177/107769908506200205>
- Hahn, D. A., & Cummins, R. G. (2014). Effects of Attractiveness, Gender, and Athlete–Reporter Congruence on Perceived Credibility of Sport Reporters. *International Journal of Sport Communication*, 7(1), 34–47. <https://doi.org/10.1123/IJSC.2013-0113>
- Hammersley, M., & Gomm, R. (1997). Bias in social research. *Sociological Research Online*, 2(1), 1–13.
- Hawley, K. (2012). *Trust: A Very Short Introduction*. Oxford University Press.
- Hayakawa, S. I. (1939). *Language in Action*. College Typing Company.
- Hayakawa, S. I. (1949). *Language in Thought and Action*. Harcourt, Brace.
- Hayes, M. H., & Patterson, D. G. (1921). Experimental development of the graphic rating method. *Psychol Bull*, 18(1), 98–9.
- Hays, R. G., & Reisner, A. E. (1991). Farm Journalists and Advertiser Influence: Pressures on Ethical Standards. *Journalism Quarterly*, 68(1–2), 172–178. <https://doi.org/10.1177/107769909106800118>
- Helman, E., Graber, E. C., Hoffman, L. H., & Gaertner, S. L. (2012). Warmth and Competence: A Content Analysis of Photographs Depicting American Presidents. *Psychology of Popular Media Culture*, 1, 46–52.
- Heider, D., McCombs, M., & Poindexter, P. M. (2005). What the Public Expects of Local News: Views on Public and Traditional Journalism. *Journalism & Mass Communication Quarterly*, 82(4), 952–967. <https://doi.org/10.1177/107769900508200412>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). The weirdest people in the world? *BEHAVIORAL AND BRAIN SCIENCES*, 33, 61–135.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences*, 33(2–3), 111–135. <https://doi.org/10.1017/S0140525X10000725>
- Hernandez, I., & Preston, J. L. (2013). Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology*, 49(1), 178–182. <https://doi.org/10.1016/j.jesp.2012.08.010>

- Heuer, C. A., McClure, K. J., & Puhl, R. M. (2011). Obesity Stigma in Online News: A Visual Content Analysis. *Journal of Health Communication, 16*(9), 976–987. <https://doi.org/10.1080/10810730.2011.561915>
- Hill, T., & Lewicki, P. (2006). *Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining*. StatSoft, Inc.
- Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing and Management, 44*, 1467–1484.
- Hoffman, L. (2013). Breaking Boundaries| Political Interviews: Examining Perceived Media Bias and Effects Across TV Entertainment Formats. *International Journal of Communication, 7*(0), 18.
- Hofstetter, C. R. (1976). *Bias in the News: Network Television Coverage of the 1972 Election Campaign*. <http://eric.ed.gov/?id=ED147814>
- Hofstetter, C. R. (1979). Perception of News Bias in 1972 Presidential Campaign. *Journalism Quarterly, 56*(2), 370–374. <https://doi.org/10.1177/107769907905600222>
- Hogarth, R. M. (1987). *Judgment and Choice: The Psychology of Decision*. Wiley.
- Holovaty, A., & Kaplan-Moss, J. (2009). *The Definitive Guide to Django: Web Development Done Right*. Apress.
- Hong, T. (2006). The influence of structural and message features on Web site credibility. *Journal of The American Society for Information Science and Technology, 57*(1), 114–127. <https://doi.org/10.1002/asi.20258>
- Hope, B. G., & Li, Z. (2003). Online Newspaper Quality Factors: The impact of Gender, Age, and Skill Level. *ACIS 2003 Proceedings, 39*.
- Horvit, B. (2006). International News Agencies and the War Debate of 2003. *International Communication Gazette, 68*(5–6), 427–447. <https://doi.org/10.1177/1748048506068722>
- Hosking, G. A. (2014). *Trust: A History*. Oxford University Press.
- Houston, J. B., Hansen, G. J., & Nisbett, G. S. (2011). Influence of User Comments on Perceptions of Media Bias and Third-Person Effect in Online News. *Electronic News, 5*(2), 79–92. <https://doi.org/10.1177/1931243111407618>
- Hovland, C. I. (1948). Social Communication. *Proceedings of the American Philosophical Society, 92*(5), 371–375. <https://doi.org/10.2307/3143048>
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; psychological studies of opinion change* (Vol. xii). Yale University Press.
- Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on Mass Communication*. (Vol. 3). Princeton University Press. <http://psycnet.apa.org/psycinfo/1950-00627-000>
- Hovland, C. I., & Weiss, W. (1951). The Influence of Source Credibility on Communication Effectiveness. *The Public Opinion Quarterly, 15*(4), 635–650. <https://doi.org/10.2307/2745952>
- Howard, P. N., & Kollanyi, B. (2016). *Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum* (SSRN Scholarly Paper ID 2798311). Social Science Research Network. <https://papers.ssrn.com/abstract=2798311>
- Hu, Y., & Sundar, S. S. (2009). Effects of Online Health Sources on Credibility and Behavioral Intentions. *Communication Research, 36*(1), 11–26. <https://doi.org/10.1177/0093650209351512>
- Huang, Z., Olteanu, A., & Aberer, K. (2013). CredibleWeb: A platform for web credibility evaluation. *CHI '13 Extended Abstracts on Human Factors in Computing Systems, 1887–1892*.
- Hughes, C. A., & Western, J. S. (1973). The geographical sources of foreign news in Australian newspapers. *Australian Outlook, 27*(1), 86–97. <https://doi.org/10.1080/10357717308444462>
- Hughes, S., & Lawson, C. H. (2004). Propaganda and Crony Capitalism: Partisan Bias in Mexican Television News. *Latin American Research Review, 39*(3), 81–105. <https://doi.org/10.1353/lar.2004.0050>
- Iyengar, S., & Hahn, K. S. (2009). Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *Journal of Communication, 59*(1), 19–39. <https://doi.org/10.1111/j.1460-2466.2008.01402.x>

- Iyyer, M., Enns, P., Boyd-Graber, J. L., & Resnik, P. (2014). Political Ideology Detection Using Recursive Neural Networks. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 1*, 1113–1122. http://www.umiacs.umd.edu/~jbg/temp/2014_RNN_framing.paper.pdf
- Jiwani, Y. (1993). *By omission and commission: "race" and representation in Canadian television news* [PhD Thesis]. Theses (Dept. of Communication)/Simon Fraser University.
- Johnson, T. J. (1993). Filling Out the Racing Form: How the Media Covered the Horse Race in the 1988 Primaries. *Journalism Quarterly, 70*(2), 300–310. <https://doi.org/10.1177/107769909307000206>
- Johnson, T. J., & Kaye, B. K. (1998). Cruising is believing?: Comparing Internet and traditional sources on media credibility measures. *Journalism & Mass Communication Quarterly, 75*(2), 325–340.
- Johnson, T. J., & Kaye, B. K. (2000). Using is believing: The influence of reliance on the credibility of online political information among politically interested Internet users. *Journalism & Mass Communication Quarterly, 77*(4), 865–879.
- Johnson, T. J., & Kaye, B. K. (2002). Webbelievability: A Path Model Examining How Convenience and Reliance Predict Online Credibility. *Journalism & Mass Communication Quarterly, 79*(3), 619–642.
- Johnson, T. J., & Kaye, B. K. (2004). Wag the blog: How reliance on traditional media and the internet influence credibility perceptions of weblogs among blog users. *Journalism & Mass Communication Quarterly, 81*(3), 622–642.
- Johnson, T. J., & Kaye, B. K. (2010). Still Cruising and Believing? An Analysis of Online Credibility Across Three Presidential Campaigns. *American Behavioral Scientist, 54*(1), 57–77. <https://doi.org/10.1177/0002764210376311>
- Jones, S. (2008). Television News: Geographic and Source Biases, 1982 - 2004. *International Journal of Communication, 2*(0), 30.
- Jung, D. J. (1996). *The Federal Communications Commission, the Broadcast Industry, and the Fairness Doctrine: 1981-1987*. UPA.
- Jung, W. S., Chung, M.-Y., & Rhee, E. S. (2017). The Effects of Attractiveness and Source Expertise on Online Health Sites. *Health Communication, 0*(0), 1–10. <https://doi.org/10.1080/10410236.2017.1323364>
- Kahn, K. F., & Kenney, P. J. (2002). The Slant of the News: How Editorial Endorsements Influence Campaign Coverage and Citizens' Views of Candidates. *American Political Science Review, 96*(2), 381–394. <https://doi.org/10.1017/S0003055402000230>
- Karimov, F. P., Brengman, M., & Van Hove, L. (2011). The Effect Of Website Design Dimensions On Initial Trust: A Synthesis Of The Empirical Literature. *Journal of Electronic Commerce Research, 12*(4).
- Keeble, R. (2008). *Ethics for Journalists*. Taylor & Francis.
- Kelman, H. C., & Hovland, C. I. (1953). "Reinstatement" of the communicator in delayed measurement of opinion change. *The Journal of Abnormal and Social Psychology, 48*(3), 327–335. <https://doi.org/10.1037/h0061861>
- Kenix, L. J. (2013). A converging image? Commercialism and the visual identity of alternative and mainstream news websites. *Journalism Studies, 14*(6), 835–856. <https://doi.org/10.1080/1461670X.2012.754239>
- Kenney, K., & Simpson, C. (1993). Was Coverage of the 1988 Presidential Race by Washington's Two Major Dailies Biased? *Journalism Quarterly, 70*(2), 345–355. <https://doi.org/10.1177/107769909307000210>
- Kepplinger, H. M. (1982). Visual Biases In Television Campaign Coverage. *Communication Research, 9*(3), 432–446. <https://doi.org/10.1177/009365082009003005>
- Kerrick, J. S., Anderson, T. E., & Swales, L. B. (1964). Balance and the Writer's Attitude in News Stories and Editorials. *Journalism Quarterly, 41*(2), 207–215. <https://doi.org/10.1177/107769906404100207>
- Kim, E., & Lowrey, W. (2018). Does Local News Side With Local Organizations? A Case Study of Boosterism and Dependence on Local and National Sources. *International Journal of Sport Communication, 11*(1), 123–140.

- Kim, Hyejeong, & Niehm, L. S. (2009). The Impact of Website Quality on Information Quality, Value, and Loyalty Intentions in Apparel Retailing. *Journal of Interactive Marketing*, 23(3), 221–233. <https://doi.org/10.1016/j.intmar.2009.04.009>
- Kim, Hyojin, Park, S.-Y., & Bozeman, I. (2011). Online health information search and evaluation: Observations and semi-structured interviews with college students and maternal health experts. *Health Information & Libraries Journal*, 28(3), 188–199. <https://doi.org/10.1111/j.1471-1842.2011.00948.x>
- Kim, J., & Moon, J. Y. (1998). Designing towards emotional usability in customer interfaces—Trustworthiness of cyber-banking system interfaces. *Interacting with Computers*, 10(1), 1–29.
- Kim, S., & Stoel, L. (2004a). Apparel retailers: Website quality dimensions and satisfaction. *Journal of Retailing and Consumer Services*, 11(2), 109–117. [https://doi.org/10.1016/S0969-6989\(03\)00010-9](https://doi.org/10.1016/S0969-6989(03)00010-9)
- Kim, S., & Stoel, L. (2004b). Dimensional hierarchy of retail website quality. *Information & Management*, 41(5), 619–633. <https://doi.org/10.1016/j.im.2003.07.002>
- Kim, Y., & Bae, J. (2006). Korean practitioners and journalists: Relational influences in news selection. *Public Relations Review*, 32(3), 241–245.
- Kingsbury, S., & Hart, H. (1933). Measuring the Ethics of American Newspapers III. Newspaper Bias on Congressional Controversies. *Journalism Bulletin*, 10(4), 323–342. <https://doi.org/10.1177/107769903301000411>
- Kiousis, S. (2001a). Public Trust or Mistrust? Perceptions of Media Credibility in the Information Age. *Mass Communication and Society*, 4(4), 381–403. https://doi.org/10.1207/S15327825MCS0404_4
- Kiousis, S. (2001b). Public Trust or Mistrust? Perceptions of Media Credibility in the Information Age. *Mass Communication and Society*, 4(4), 381–403. https://doi.org/10.1207/S15327825MCS0404_4
- Kishen, K. (1942). On Latin and Hyper-Graeco-Latin Cubes and Hyper Cubes. *Current Science*, 11(3), 98–99.
- Kittur, A., Chi, E. H., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 453–456. <https://doi.org/10.1145/1357054.1357127>
- Klein, M. W., & Maccoby, N. (1954). Newspaper Objectivity in the 1952 Campaign. *Journalism Bulletin*, 31(3), 285–296. <https://doi.org/10.1177/107769905403100301>
- Kline, S. (1979). Creating Reality: How Television News Distorts Events. *Canadian Journal of Communication*, 5(3). <https://doi.org/10.22230/cjc.1979v5n3a208>
- Kline, S. (1981). National Perspective and News Bias: A comparison of national news broadcasts. *Canadian Journal of Communication*, 7(4).
- Klinenberg, E. (2005). Convergence: News Production in a Digital Age. *The Annals of the American Academy of Political and Social Science*, 597(1), 48–64. <https://doi.org/10.1177/0002716204270346>
- Kobre, S. (1953). How Florida Dailies Handled the 1952 Presidential Campaign. *Journalism Bulletin*, 30(2), 163–169. <https://doi.org/10.1177/107769905303000203>
- Koffka, K. (1935). *Principles of Gestalt psychology*. Harcourt, Brace.
- Koufaris, M., & Hampton-Sosa, W. (2004). The development of initial trust in an online company by new customers. *Information & Management*, 41(3), 377–397. <https://doi.org/10.1016/j.im.2003.08.004>
- Krestel, R., Wall, A., & Nejd, W. (2012). Treehugger or petrolhead?: Identifying bias by comparing online news articles with political speeches. *Proceedings of the 21st International Conference Companion on World Wide Web*, 547–548. <https://doi.org/10.1145/2187980.2188120>
- Kriesberg, M. (1946). Soviet News in the “New York Times.” *Public Opinion Quarterly*, 10(4), 540–564. <https://doi.org/10.1093/poq/10.4.540>
- Kuypers, J. A. (2002). *Press Bias and Politics: How the Media Frame Controversial Issues*. Greenwood Publishing Group.

- Kwak, H., & An, J. (2016). Revealing the Hidden Patterns of News Photos: Analysis of Millions of News Photos Using GDELT and Deep Learning-based Vision APIs. *Tenth International AAAI Conference on Web and Social Media*. <http://arxiv.org/abs/1603.04531>
- Ladhari, R., & Michaud, M. (2015). EWOM effects on hotel booking intentions, attitudes, trust, and website perceptions. *International Journal of Hospitality Management*, 46, 36–45. <https://doi.org/10.1016/j.ijhm.2015.01.010>
- Lagun, D., & Lalmas, M. (2016). Understanding User Attention and Engagement in Online News Reading. *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, 113–122. <https://doi.org/10.1145/2835776.2835833>
- Larcinese, V., Puglisi, R., & Snyder Jr, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers. *Journal of Public Economics*, 95(9), 1178–1189.
- Larson, M. S. (1989). Presidential News Coverage and "All Things Considered": National Public Radio and News Bias. *Presidential Studies Quarterly*, 19(2), 347–353.
- Latham, O. (2013). *Bias at the Beeb?: A Quantitative Study of Slant in BBC Online Reporting: Summary* (p. 13) [Think Tank Report]. Centre for Policy Studies. <https://www.cps.org.uk/research/bias-at-the-beeb/#>
- Lee, A. S. (1991). Integrating Positivist and Interpretive Approaches to Organizational Research. *Organization Science*, 2(4), 342–365. <https://doi.org/10.1287/orsc.2.4.342>
- Lee, T.-T. (2008). Bias in the News. In W. Donsbach (Ed.), *The International Encyclopedia of Communication*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781405186407.wbiecb014>
- Lee, T.-T. (2010). Why They Don't Trust the Media: An Examination of Factors Predicting Trust. *American Behavioral Scientist*, 54(1), 8–21. <https://doi.org/10.1177/0002764210376308>
- Lee, & Ware, B. (2002). *Open Source Development with LAMP: Using Linux, Apache, MySQL and PHP*. Addison-Wesley Longman Publishing Co., Inc.
- Leetaru, K. (2019, January 22). The Media's BuzzFeeding Frenzy: Tracking the Cohen Story | RealClearPolitics. *RealClearPolitics*. https://www.realclearpolitics.com/articles/2019/01/22/the_medias_buzzfeeding_frenzy_tracking_the_cohen_story_139237.html
- Lefever, E. W. (1974). *TV and national defense: An analysis of CBS news, 1972-73* (1st edition). Institute for American Strategy Press.
- Lehane, M. (2018, March 2). Govt Strategic Communications Unit to be reviewed. *RTE.Ie*. <https://www.rte.ie/news/politics/2018/0302/944598-government-strategic-comms-unit/>
- Lengauer, G., Esser, F., & Berganza, R. (2012). Negativity in political news: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 179–202. <https://doi.org/10.1177/1464884911427800>
- Lewis, J. D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces*, 63(4), 967–985. <https://doi.org/10.1093/sf/63.4.967>
- Li, F., Pieńkowski, D., van Moorsel, A., & Smith, C. (2011). A Holistic Framework for Trust in Online Transactions. *International Journal of Management Reviews*, 14(1), 85–103. <https://doi.org/10.1111/j.1468-2370.2011.00311.x>
- Li, H., & Tang, L. (2009). The representation of the Chinese product crisis in national and local newspapers in the United States. *Public Relations Review*, 35(3), 219–225. <https://doi.org/10.1016/j.pubrev.2009.05.007>
- Lin, H.-F. (2007). The Impact of Website Quality Dimensions on Customer Satisfaction in the B2C E-commerce Context. *Total Quality Management & Business Excellence*, 18(4), 363–378. <https://doi.org/10.1080/14783360701231302>
- Lin, H.-F. (2011). An empirical investigation of mobile banking adoption: The effect of innovation attributes and knowledge-based trust. *International Journal of Information Management*, 31(3), 252–260. <https://doi.org/10.1016/j.ijinfomgt.2010.07.006>

- Lin, W.-Y., Lo, V., & Wang, T.-L. (2011). Bias in television foreign news in China, Hong Kong, and Taiwan. *Chinese Journal of Communication, 4*(3), 293–310. <https://doi.org/10.1080/17544750.2011.594555>
- Lin, Y.-R., Bagrow, J. P., & Lazer, D. (2011). More Voices Than Ever? Quantifying Media Bias in Networks. *ICWSM - Fifth International AAAI Conference on Weblogs and Social Media, 1*(arXiv: 1111.1227), 1.
- Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology, 25*(2), 115–126. <https://doi.org/10.1080/01449290500330448>
- Liu, S.-D. (2006). China's popular nationalism on the internet. Report on the 2005 anti-Japan network struggles. *Inter-Asia Cultural Studies, 7*(1), 144–155.
- Lo, V. (1995). Journalists' bias in news sources selection. *Mass Communication Research, 50*, 1–13.
- Lo, V., King, P., Chen, C., & Huang, H. (1996). Political bias in the news coverage of Taiwan's first presidential election: A comparative analysis of broadcast TV and cable TV news. *Asian Journal of Communication, 6*(2), 43–64. <https://doi.org/10.1080/01292989609364743>
- Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2002). WebQual: A measure of website quality. *Marketing Theory and Applications, 13*(3), 432–438.
- Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2007). WebQual: An Instrument for Consumer Evaluation of Web Sites. *International Journal of Electronic Commerce, 11*(3), 51–87. <https://doi.org/10.2753/JEC1086-4415110302>
- Long, M. M., & Chiagouris, L. (2006). The role of credibility in shaping attitudes toward nonprofit websites. *International Journal of Nonprofit and Voluntary Sector Marketing, 11*(3), 239–249. <https://doi.org/10.1002/nvsm.36>
- Lott, J. R., & Hassett, K. A. (2014). Is newspaper coverage of economic events politically biased? *Public Choice, 160*(1–2), 65–108. <https://doi.org/10.1007/s11127-014-0171-5>
- Lowrey, W. (2002). The influential designer: Explaining variability in control over news presentation work. *Visual Communication Quarterly, 9*(1), 4–13. <https://doi.org/10.1080/15551390209363473>
- Lowry, D. T. (1973). Measures of Network News Bias in the 1972 Presidential Campaign. *Journal of Broadcasting, 18*, 387.
- Lowry, D. T. (1986). Establishing Construct Validity of the Hayakawa-Lowry News Bias Categories. *Journalism & Mass Communication Quarterly, 63*(3), 573–580. <https://doi.org/10.1177/107769908606300319>
- Lowry, D. T. (2008). Network TV News Framing of Good Vs. Bad Economic News under Democrat and Republican Presidents: A Lexical Analysis of Political Bias. *Journalism & Mass Communication Quarterly, 85*(3), 483–498. <https://doi.org/10.1177/107769900808500301>
- Lowry, D. T., & Shidler, J. A. (1998). The Sound Bites, the Biters, and the Bitten: A Two-Campaign Test of the Anti-Incumbent Bias Hypothesis in Network TV News. *Journalism & Mass Communication Quarterly, 75*(4), 719–729. <https://doi.org/10.1177/107769909807500407>
- Lowry, P. B., Wilson, D. W., & Haig, W. L. (2014). A Picture is Worth a Thousand Words: Source Credibility Theory Applied to Logo and Website Design for Heightened Credibility and Consumer Trust. *International Journal of Human-Computer Interaction, 30*(1), 63–93. <https://doi.org/10.1080/10447318.2013.839899>
- Lu, K. (2017). *Pew Research Center, Growth in mobile news use driven by older adults*. Pew Research Center. <http://www.pewresearch.org/fact-tank/2017/06/12/growth-in-mobile-news-use-driven-by-older-adults/>
- Lucassen, T., Muilwijk, R., Noordzij, M. L., & Schraagen, J. M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of The American Society for Information Science and Technology, 64*(2), 254–264. <https://doi.org/10.1002/asi.22743>
- Ma, H., Liu, X., & Shen, Z. (2016). User Fatigue in Online News Recommendation. *Proceedings of the 25th International Conference on World Wide Web, 1363–1372*. <https://doi.org/10.1145/2872427.2874813>
- Maass, A., Salvi, D., Arcuri, L., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology, 57*(6), 981–993.

- Maass, Anne. (1999). Linguistic Intergroup Bias: Stereotype Perpetuation Through Language. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 31, pp. 79–121). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60272-5](https://doi.org/10.1016/S0065-2601(08)60272-5)
- MacGill Hughes, H. (1938). Review of Newspaper and the News [Review of *Review of Newspaper and the News*, by S. M. Kingsbury & H. Hart]. *American Journal of Sociology*, 43(4), 669–671.
- Mantler, G., & Whiteman, D. (1995). Attention to Candidates and Issues in Newspaper Coverage of 1992 Presidential Campaign. *Newspaper Research Journal*, 16(3), 14–28. <https://doi.org/10.1177/073953299501600302>
- Markham, D. (1965). *The Dimensions of Source Credibility of Television Newscasters*. <https://academic.oup.com/joc/article/18/1/57-64/4560884>
- Markham, D. (1968). The Dimensions of Source Credibility of Television Newscasters. *Journal of Communication*, 18(1), 57–64. <https://doi.org/10.1111/j.1460-2466.1968.tb00055.x>
- Martin, S. R. (1988). Proximity of Event as Factor in Selection of News Sources. *Journalism Quarterly*, 65(4), 986–989. <https://doi.org/10.1177/107769908806500424>
- Matejka, J., Glueck, M., Grossman, T., & Fitzmaurice, G. (2016). *The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales*. 5421–5432. <https://doi.org/10.1145/2858036.2858063>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>
- McCain, T. A., Chilberg, J., & Wakshlag, J. (1977). The effect of camera angle on source credibility and attraction. *Journal of Broadcasting*, 21(1), 35–46. <https://doi.org/10.1080/08838157709363815>
- McCarthy, J. D., & McPhail, C. (1997). The institutionalization of protest in the United States. In D. S. Meyer & S. Tarrow (Eds.), *The Social Movement Society: Contentious Politics for a New Century* (pp. 83–110). Rowman & Littlefield Publishers.
- McCarthy, J. D., McPhail, C., & Smith, J. (1996). Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991. *American Sociological Review*, 61(3), 478–499. <https://doi.org/10.2307/2096360>
- McCarthy, J. D., McPhail, C., Smith, J., & Crishock, L. (1999). Electronic and Print Media Representations of Washington D.C. Demonstrations, 1982 and 1991. In D. Rucht, R. Koopmans, & F. Neidhardt (Eds.), *Acts of Dissent: New Developments in the Study of Protest* (pp. 113–130). Rowman & Littlefield. <http://d-scholarship.pitt.edu/20703/>
- McChesney, R. W. (2003). Theses on media deregulation. *Media, Culture & Society*, 25(1), 125–133.
- McCluskey, M. (2005). A Content Analysis Of 2004 Presidential Election Headlines Of The Los Angeles Times And The Washington Times. *Electronic Theses and Dissertations*. <http://stars.library.ucf.edu/etd/358>
- McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176–187.
- McCormick, I. A., Walkey, F. H., & Green, D. E. (1986). Comparative perceptions of driver ability—A confirmation and expansion. *Accident Analysis & Prevention*, 18(3), 205–208. [https://doi.org/10.1016/0001-4575\(86\)90004-7](https://doi.org/10.1016/0001-4575(86)90004-7)
- McCroskey, J. C. (1966). Scales for the measurement of ethos. *Speech Monographs*, 33(1), 65–72. <https://doi.org/10.1080/03637756609375482>
- McCroskey, J. C., & Jenson, T. A. (1975). Image of mass media news sources. *Journal of Broadcasting*, 19(2), 169–180. <https://doi.org/10.1080/08838157509363777>
- McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, 66(1), 90–103. <https://doi.org/10.1080/03637759909376464>
- McCroskey, J. C., & Young, T. J. (1981). Ethos and credibility: The construct and its measurement after three decades. *Central States Speech Journal*, 32(1), 24–34. <https://doi.org/10.1080/10510978109368075>

- McEnroe, J., Loughlin, E., & McConnell, D. (2018, March 6). Leo Varadkar all but confirms the demise of Strategic Communications Unit | Irish Examiner. *Irish Examiner*.
<https://www.irishexaminer.com/breakingnews/ireland/leo-varadkar-all-but-confirms-the-demise-of-strategic-communications-unit-831068.html>
- McGrath, K., & Gaziano, C. (1986). Dimensions of Media Credibility: Highlights of the 1985 ASNE Survey. *Newspaper Research Journal*, 7(2), 55–67. <https://doi.org/10.1177/073953298600700207>
- McKnight, D. H., & Chervany, N. L. (1996). *The meanings of trust*. Carlson School of Management, Univ. of Minnesota.
- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, 13(3), 334–359. <https://doi.org/10.1287/isre.13.3.334.81>
- McKnight, D. H., & Kacmar, C. J. (2007). Factors and effects of information credibility. *Proceedings of the Ninth International Conference on Electronic Commerce*, 423. <https://doi.org/10.1145/1282100.1282180>
- McLeod, D. M., & Detenber, B. H. (1999). Framing Effects of Television News Coverage of Social Protest. *Journal of Communication*, 49(3), 3–23. <https://doi.org/10.1111/j.1460-2466.1999.tb02802.x>
- McMillan, K. (2005). Racial Discrimination and Political Bias on Talkback Radio in New Zealand: Assessing the Evidence. *Political Science*, 57(2), 75–91. <https://doi.org/10.1177/003231870505700208>
- McPhail, C., & Schweingruber, D. (1998). Unpacking Protest Events: A Description Bias Analysis of Media Records with Systematic Direct Observations of Collective Action—The 1995 March for Life. In D. Rucht, R. Koopmans, & F. Neidhardt (Eds.), *Acts of dissent: New developments in the study of protest* (pp. 164–195). Rowman & Littlefield.
- McQuail, D. (1992). *Media Performance: Mass Communication and the Public Interest*. SAGE.
- Meadors, J. D., & Murray, C. B. (2014). Measuring Nonverbal Bias Through Body Language Responses to Stereotypes. *Journal of Nonverbal Behavior*, 38(2), 209–229. <https://doi.org/10.1007/s10919-013-0172-y>
- Merron, J., & Gaddy, G. D. (1986). Editorial Endorsements and News Play: Bias in Coverage of Ferraro’s Finances. *Journalism Quarterly*, 63(1), 127–137. <https://doi.org/10.1177/107769908606300119>
- Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of The American Society for Information Science and Technology*, 58(13), 2078–2091. <https://doi.org/10.1002/asi.20672>
- Metzger, M. J., & Flanagin, A. J. (2015). Psychological Approaches to Credibility Assessment Online. *The Handbook of the Psychology of Communication Technology*, 32, 445–466.
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. (2003). *Credibility in the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment.pdf*. <http://www.comm.ucsb.edu/faculty/flanagin/CV/MetzgerFlanaginetal2003%28CY%29.pdf>
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & Mccann, R. M. (2003). Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment. In P. Kalbfleisch (Ed.), *Communication Yearbook* (Vol. 27, pp. 293–335). Lawrence Erlbaum Associates. <https://www.tandfonline.com/doi/full/10.1080/23808985.2003.11679029>
- Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41(3), 271–290. [https://doi.org/10.1016/S0360-1315\(03\)00049-6](https://doi.org/10.1016/S0360-1315(03)00049-6)
- Meyer, P. (1988). Defining and measuring credibility of newspapers: Developing an index. *Journalism Quarterly*, 65, 567.
- Meyer, T. P. (1972). News reporter bias: A case study in selective perception. *Journal of Broadcasting*, 16(2), 195–203. <https://doi.org/10.1080/08838157209386343>
- Miller, A., Coleman, R., & Granberg, D. (2007). TV Anchors, Elections & Bias: A Longitudinal Study of the Facial Expressions of Brokaw Rather Jennings. *Visual Communication Quarterly*, 14(4), 244–257. <https://doi.org/10.1080/15551390701730232>

- Mitchell, A., Gottfried, J., Kiley, J., & Matsa, K. E. (2014, October 21). Political Polarization & Media Habits. *Pew Research Center's Journalism Project*. <http://www.journalism.org/2014/10/21/political-polarization-media-habits/>
- Mohd-Any, A. A., Winklhofer, H., & Ennew, C. (2015). Measuring users' value experience on a travel website (e-value) what value is cocreated by the user? *Journal of Travel Research*, 54(4), 496–510.
- Montiel, C. J., Salvador, A. M. O., See, D. C., & De Leon, M. M. (2014). Nationalism in local media during international conflict: Text mining domestic news reports of the China–Philippines maritime dispute. *Journal of Language and Social Psychology*, 33(5), 445–464.
- Moriarty, S. E., & Garramone, G. M. (1986). A Study of Newsmagazine Photographs of the 1984 Presidential Campaign. *Journalism Quarterly*, 63(4), 728–734. <https://doi.org/10.1177/107769908606300408>
- Moriarty, S. E., & Popovich, M. N. (1991). Newsmagazine Visuals and the 1988 Presidential Election. *Journalism Quarterly*, 68(3), 371–380. <https://doi.org/10.1177/107769909106800307>
- Morris, J. S. (2005). The Fox News Factor. *Harvard International Journal of Press/Politics*, 10(3), 56–79. <https://doi.org/10.1177/1081180X05279264>
- Moss, P. (1988). Words, Words, Words: Radio News Discourses and How They Work. *European Journal of Communication*, 3(2), 207–230. <https://doi.org/10.1177/0267323188003002006>
- Mrogers, E., & Wdearing, J. (1988). Agenda-Setting Research: Where Has It Been, Where Is It Going? *Annals of the International Communication Association*, 11(1), 555–594. <https://doi.org/10.1080/23808985.1988.11678708>
- Muddiman, A., & Stroud, N. J. (2017). News Values, Cognitive Biases, and Partisan Incivility in Comment Sections: Uncivil Comments. *Journal of Communication*, 67(4), 586–609. <https://doi.org/10.1111/jcom.12312>
- Mughan, A. (1996). Television can matter: Bias in the 1992 general election. *British Elections and Parties Yearbook*, 6(1), 128–142. <https://doi.org/10.1080/13689889608412976>
- Mukherjee, A., & Nath, P. (2003). A model of trust in online relationship banking. *International Journal of Bank Marketing*, 21(1), 5–15. <https://doi.org/10.1108/02652320310457767>
- Mullainathan, S., & Shleifer, A. (2005). The Market for News. *The American Economic Review*, 95(4), 1031–1053.
- Mullen, B., Futrell, D., Stairs, D., Tice, D. M., Baumeister, R. F., Dawson, K. E., Riordan, C. A., Radloff, C. E., Goethals, G. R., & Kennedy, J. G. (1986). Newscasters' facial expressions and voting behavior of viewers: Can a smile elect a president? *Journal of Personality and Social Psychology*, 51(2), 291.
- Murphy, L. W. (1934). Notes on the Kingsbury-Hart Study. *Journalism Bulletin*, 11(4), 382–391. <https://doi.org/10.1177/107769903401100404>
- Muscati, S. A. (2002). Arab/Muslim “Otherness”: The Role of Racial Constructions in the Gulf War and the Continuing Crisis with Iraq. *Journal of Muslim Minority Affairs*, 22(1), 131–148. <https://doi.org/10.1080/13602000220124872>
- Narwal, V., Salih, M. H., Lopez, J. A., Ortega, A., O'Donovan, J., Höllerer, T., & Savage, S. (2017). Automated Assistants to Identify and Prompt Action on Visual News Bias. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2796–2801. <https://doi.org/10.1145/3027063.3053227>
- Navia, P., & Osorio, R. (2015). El Mercurio Lies, and La Tercera Lies More. Political Bias in Newspaper Headlines in Chile, 1994–2010. *Bulletin of Latin American Research*, 34(4), 467–485. <https://doi.org/10.1111/blar.12364>
- Nazir, A. K., Zafar, I., Shaheen, A., Maqbool, B., & Qamar, U. (2018). GGSE-Website Usability Evaluation Framework. *Science and Information Conference*, 426–436.
- Nelson, M. J. (2013). Statistical Power and Effect Size in Informative Retrieval Experiments. *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*, 0(0). <https://www.cais-acsi.ca/ojs/index.php/cais/article/view/523>
- Newhagen, J. E., & Reeves, B. (1992). The Evening's Bad News: Effects of Compelling Negative Television News Images on Memory. *Journal of Communication*, 42(2), 25–41. <https://doi.org/10.1111/j.1460-2466.1992.tb00776.x>

- Newhagen, J., & Nass, C. (1989). Differential criteria for evaluating credibility of newspapers and TV news. *Journalism and Mass Communication Quarterly*, 66(2), 277. <http://search.proquest.com/openview/5a86f9e4ca056a9cadada145bf3beb9b/1?pq-origsite=gscholar>
- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175–220.
- Niven, D. (1999). Partisan Bias in the Media? A New Test. *Social Science Quarterly*, 80(4), 847–857.
- Niven, D. (2002). *Tilt?: The Search for Media Bias*. Greenwood Publishing Group.
- Noel Whelan. (2018, March 9). Noel Whelan: Strategic Communications Unit a good idea badly executed. *The Irish Times*. <https://www.irishtimes.com/opinion/noel-whelan-strategic-communications-unit-a-good-idea-badly-executed-1.3419922>
- Nokelainen, T., & Kanninen, J. (2018). Coverage bias in business news: Evidence and methodological implications. *Management Research Review*, 41(4), 487–503. <https://doi.org/10.1108/MRR-02-2017-0048>
- Nord, D. P. (1985). The ideology of the press: Why press bias isn't what it appears to be. *The Cresset*, 49, 14–17.
- Ó Cionnaith, F., McEnroe, J., & Leogue, J. (2018, March 8). Leo Varadkar faces PAC grilling over Strategic Communications Unit | Irish Examiner. *Irish Examiner*. <https://www.irishexaminer.com/ireland/leo-varadkar-faces-pac-grilling-over-strategic-communications-unit-467983.html>
- Ognianova, E. (1998). *The value of journalistic identity on the World Wide Web*. Mass Communication and Society Division, Association for Education in Journalism and Mass Communication, Baltimore.
- O'Keefe, D. J. (2015). *Persuasion: Theory and Research*. SAGE Publications.
- Oliver, P. E., & Maney, G. M. (2000). Political Processes and Local Newspaper Coverage of Protest Events: From Selection Bias to Triadic Interactions. *American Journal of Sociology*, 106(2), 463–505. <https://doi.org/10.1086/316964>
- O'Neal, G. S., & Lapitsky, M. (1991). Effects of Clothing as Nonverbal Communication on Credibility of the Message Source. *Clothing and Textiles Research Journal*, 9(3), 28–34. <https://doi.org/10.1177/0887302X9100900305>
- Orellana-Rodriguez, C., & Keane, M. T. (2018). Attention to news and its dissemination on Twitter: A survey. *Computer Science Review*, 29, 74–94. <https://doi.org/10.1016/j.cosrev.2018.07.001>
- Ou, C. X., & Sia, C. L. (2010). Consumer trust and distrust: An issue of website design. *International Journal of Human-Computer Studies*, 68(12), 913–934. <https://doi.org/10.1016/j.ijhcs.2010.08.003>
- Owens, L., & Palmer, L. K. (2003). Making the News: Anarchist Counter-Public Relations on the World Wide Web. *Critical Studies in Media Communication*, 20(4), 335–361. <https://doi.org/10.1080/0739318032000142007>
- Page, B. I., & Shapiro, R. Y. (2010). *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. University of Chicago Press.
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*. <https://doi.org/10.1016/j.jbef.2017.12.004>
- Papaioannou, T. G., Ranvier, J.-E., Olteanu, A., & Aberer, K. (2012). A Decentralized Recommender System for Effective Web Credibility Assessment. *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, 704–713. <https://doi.org/10.1145/2396761.2396851>
- Park, S., Kang, S., Chung, S., & Song, J. (2012). A Computational Framework for Media Bias Mitigation. *ACM Trans. Interact. Intell. Syst.*, 2(2), 8:1–8:32. <https://doi.org/10.1145/2209310.2209311>
- Park, S., Kang, S., Chung, S., & Song, J. (2009). NewsCube: Delivering multiple aspects of news to mitigate media bias. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 443–452. <https://doi.org/10.1145/1518701.1518772>
- Park, S., Ko, M., Kim, J., Choi, H., & Song, J. (2011). NewsCube2. 0: An Exploratory Design of a Social News Website for Media Bias Mitigation. *Workshop on Social Recommender Systems*. <http://nclab.kaist.ac.kr/papers/Conference/NC2.pdf>

- Park, S., Ko, M., Liu, Y., Jin, D. Y., & Song, J. (2011). *Improving Journalism through the Web: Framework for Media Bias Mitigation*. 1–3. http://www.websci11.org/fileadmin/websci/Posters/88_paper.pdf
- Patterson, T. E., & Donsbagh, W. (1996). News decisions: Journalists as partisan actors. *Political Communication*, 13(4), 455–468. <https://doi.org/10.1080/10584609.1996.9963131>
- Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7, 101–134. <https://doi.org/10.1080/10864415.2003.11044275>
- Pelet, J. É., & Papadopoulou, P. (2012). The Effect of E-Commerce Websites' Colors on Customer Trust. In L. In (Ed.), *Mobile Applications and Knowledge Advancements in E-Business* (pp. 167–185). Idea Group Inc (IGI).
- Pennbridge, J., Moya, R., & Rodrigues, L. (1999). Questionnaire survey of California consumers' use and rating of sources of health care information including the Internet. *Western Journal of Medicine*, 171(5–6), 302–305.
- Persily, N. (2017). Can Democracy Survive the Internet? *Journal of Democracy*, 28(2), 63–76. <https://doi.org/10.1353/jod.2017.0025>
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46(1), 69.
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In *Communication and Persuasion Central and Peripheral Routes to Attitude Change* (pp. 1–24). Springer New York. http://link.springer.com/chapter/10.1007/978-1-4612-4964-1_1
- Pew. (2015a, April 1). U.S. Smartphone Use in 2015. *Pew Research Center: Internet, Science & Tech*. <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>
- Pew. (2015b, May 6). U.S. voter turnout trails most developed countries. *Pew Research Center*. <http://www.pewresearch.org/fact-tank/2015/05/06/u-s-voter-turnout-trails-most-developed-countries/>
- Pew. (2015c, June 26). Americans' Internet Access: 2000-2015. *Pew Research Center: Internet, Science & Tech*. <http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/>
- Pew, P. (2012, February 7). Pew Research Center—Cable Leads the Pack as Campaign News Source. *Pew Research Center for the People and the Press*. <http://www.people-press.org/2012/02/07/cable-leads-the-pack-as-campaign-news-source/>
- Pfau, M., Haigh, M., Gettle, M., Donnelly, M., Scott, G., Warr, D., & Wittenberg, E. (2004). Embedding Journalists in Military Combat Units: Impact on Newspaper Story Frames and Tone. *Journalism & Mass Communication Quarterly*, 81(1), 74–88. <https://doi.org/10.1177/107769900408100106>
- Pfau, M., Houston, J. B., & Semmler, S. M. (2007). *Mediating the Vote: The Changing Media Landscape in U.S. Presidential Campaigns*. Rowman & Littlefield.
- Philo, G., Beharrell, P., & Hewitt, J. (1977). One Dimensional New-Television and the Control of Explanation. In G. Philo & G. Philo (Eds.), *Trade Unions and the Media* (pp. 1–22). Springer.
- Pintak, L. (2014). Islam, identity and professional values: A study of journalists in three Muslim-majority regions. *Journalism*, 15(4), 482–503. <https://doi.org/10.1177/1464884913490269>
- Pirolli, P., Wollny, E., & Suh, B. (2009). So you know you're getting the best possible information: A tool that increases Wikipedia credibility. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1505–1508. <https://doi.org/10.1145/1518701.1518929>
- Price, C. J. (2003). Interfering Owners or Meddling Advertisers: How Network Television News Correspondents Feel About Ownership and Advertiser Influence on News Stories. *Journal of Media Economics*, 16(3), 175–188. https://doi.org/10.1207/S15327736ME1603_3
- Pride, R. A., & Wamsley, G. L. (1972). Symbol Analysis of Network Coverage of Laos Incursion. *Journalism Quarterly*, 49(4), 635–647. <https://doi.org/10.1177/107769907204900401>

- Pronin, E., Lin, D. Y., & Ross, L. (2002). The Bias Blind Spot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin*, 28(3), 369–381. <https://doi.org/10.1177/0146167202286008>
- Puglisi, R., & Snyder, J. M. (2011). Newspaper Coverage of Political Scandals. *The Journal of Politics*, 73(3), 931–950. <https://doi.org/10.1017/S0022381611000569>
- Rada, J. A. (1996). Color blind-sided: Racial bias in network television's coverage of professional football games. *Howard Journal of Communications*, 7(3), 231–239. <https://doi.org/10.1080/10646179609361727>
- Rafalak, M., Bilski, P., & Wierzbicki, A. (2014). Analysis of Demographical Factors' Influence on Websites' Credibility Evaluation. *Human-Computer Interaction: Applications and Services*, 57–68. https://doi.org/10.1007/978-3-319-07227-2_7
- Rains, S. A., & Karmikel, C. D. (2009). Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. *Computers in Human Behavior*, 25(2), 544–553. <https://doi.org/10.1016/j.chb.2008.11.005>
- Rainville, R. E., & McCormick, E. (1977). Extent of Covert Racial Prejudice in Pro Football Announcers' Speech. *Journalism Quarterly*, 54(1), 20–26. <https://doi.org/10.1177/107769907705400104>
- Ranly, D. (1979). How Religion Editors of Newspapers View their Jobs and Religion. *Journalism Quarterly*, 56(4), 844–850. <https://doi.org/10.1177/107769907905600421>
- Recasens, M., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2013). Linguistic Models for Analyzing and Detecting Biased Language. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 1, 1650–1659. http://www.cs.cornell.edu/~cristian/Biased_language_files/neutralty.pdf
- Reuter, J., & Zitzewitz, E. (2006). Do Ads Influence Editors? Advertising and Bias in the Financial Media. *The Quarterly Journal of Economics*, 121(1), 197–227. <https://doi.org/10.1093/qje/121.1.197>
- Riegelsberger, J., Sasse, M. A., & McCarthy, J. D. (2005). The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies*, 62(3), 381–422. <https://doi.org/10.1016/j.ijhcs.2005.01.001>
- Riegelsberger, J., Sasse, M. A., & McCarthy, J. D. (2003). Shiny Happy People Building Trust?: Photos on e-Commerce Websites and Consumer Trust. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 121–128. <https://doi.org/10.1145/642611.642634>
- Rieh, S. Y. (2002). Judgment of Information Quality and Cognitive Authority in the Web. *Journal of The American Society for Information Science and Technology*, 53(2), 145–161.
- Rieh, S. Y., & Belkin, N. (2000). Interaction on the Web: Scholars' judgement of information quality and cognitive authority. *Proceedings of the Annual Meeting of the American Society for Information Science*, 37, 25–38. http://rieh.people.si.umich.edu/~rieh/papers/rieh_asis2000.pdf
- Rieh, S. Y., & Belkin, N. J. (1998). Understanding judgment of information quality and cognitive authority in the WWW. *Proceedings of the 61st Annual Meeting of the American Society for Information Science*, 35, 279–289. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.8991&rep=rep1&type=pdf>
- Rieh, S. Y., & Hilligoss, B. (2008). College students' credibility judgments in the information-seeking process. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 49–72). http://rieh.people.si.umich.edu/papers/rieh_credibility_chapter3.pdf
- Rinallo, D., & Basuroy, S. (2009). Does advertising spending influence media coverage of the advertiser? *Journal of Marketing*, 73(6), 33–46.
- Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing and Management*, 44(1), 386–399. <https://doi.org/10.1016/j.ipm.2007.02.003>
- Robins, D., Holmes, J., & Stansbury, M. (2010). Consumer health information on the Web: The relationship of visual design and perceptions of credibility. *Journal of The American Society for Information Science and Technology*, 61(1), 13–29. <https://doi.org/10.1002/asi.21224>

- Robinson, C. (2016). Visual Press Bias In A Multi-Party Electoral Context. *Design & Democracy*, 37.
- Robinson, M. J., & Sheehan, M. A. (1983). *Over the Wire and on TV: CBS and UPI in Campaign '80*. Russell Sage Foundation.
- Rocha, Á. (2012). Framework for a global quality evaluation of a website. *Online Information Review*, 36(3), 374–382. <https://doi.org/10.1108/14684521211241404>
- Romer, D., Jamieson, K. H., & De Coteau, N. J. (1998). The Treatment of Persons of Color in Local Television News: Ethnic Blame Discourse or Realistic Group Conflict? *Communication Research*, 25(3), 286–305. <https://doi.org/10.1177/009365098025003002>
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3), 279–301. [https://doi.org/10.1016/0022-1031\(77\)90049-X](https://doi.org/10.1016/0022-1031(77)90049-X)
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not So Different After All: A Cross-Discipline View Of Trust. *Academy of Management Review*, 23(3), 393–404. <https://doi.org/10.5465/AMR.1998.926617>
- Roy, M. C., Dewit, O., & Aubert, B. A. (2001). The impact of interface usability on trust in Web retailers. *Internet Research*, 11(5), 388–398. <https://doi.org/10.1108/10662240110410165>
- Royal Commission on the Press, 1947-1949: Report*. (1949). H.M. Stationery Office.
- Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, 5(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2
- Rucinski, D. (1992). Personalized Bias in News: The Potency of the Particular? *Communication Research*, 19(1), 91–108. <https://doi.org/10.1177/009365092019001004>
- Russo, F. D. (1971). A Study of Bias in TV Coverage of The Vietnam War: 1969 And 1970. *Public Opinion Quarterly*, 35(4), 539–543. <https://doi.org/10.1086/267949>
- Sabo, D., Jansen, S. C., Tate, D., Duncan, M. C., & Leggett, S. (1996). Televising International Sport: Race, Ethnicity, And Nationalistic Bias. *Journal of Sport and Social Issues*, 20(1), 7–21. <https://doi.org/10.1177/019372396020001002>
- Sachsman, D. B. (1970). A Test of ‘Loading’: New Measure of Bias. *Journalism Quarterly*, 47(4), 759–762. <https://doi.org/10.1177/107769907004700415>
- Saez-Trumper, D., Castillo, C., & Lalmas, M. (2013). Social media news communities: Gatekeeping, coverage, and statement bias. *Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management*, 1679–1684. <https://doi.org/10.1145/2505515.2505623>
- Saks, J., Compton, J. L., Hopkins, A., & Damanhoury, K. E. (2016). Dialed In: Continuous Response Measures in Televised Political Debates and Their Effect on Viewers. *Journal of Broadcasting & Electronic Media*, 60(2), 231–247. <https://doi.org/10.1080/08838151.2016.1164164>
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk & Uncertainty*, 1(1), 7–59. <https://doi.org/10.1007/BF00055564>
- Sauro, J. (2015). SUPR-Q: A comprehensive measure of the quality of the website user experience. *Journal of Usability Studies*, 10(2).
- Schiffer, A. J. (2006). Assessing Partisan Bias in Political News: The Case(s) of Local Senate Election Coverage. *Political Communication*, 23(1), 23–39. <https://doi.org/10.1080/10584600500476981>
- Schill, D. (2012). The Visual Image and the Political Image: A Review of Visual Communication Research in the Field of Political Communication. *Review of Communication*, 12(2), 118–142. <https://doi.org/10.1080/15358593.2011.653504>
- Schindler, J., Krämer, B., & Müller, P. (2017). Looking left or looking right? Effects of newspaper layout style on the perception of political news. *European Journal of Communication*, 32(4), 348–366. <https://doi.org/10.1177/0267323117718463>

- Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66. <https://doi.org/10.1037/0033-295X.84.1.1>
- Schwarz, J., & Morris, M. (2011). Augmenting Web Pages and Search Results to Support Credibility Assessment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1245–1254. <https://doi.org/10.1145/1978942.1979127>
- Schweiger, W. (1998). Wer glaubt dem World Wide Web? In *Online-Kommunikation* (pp. 123–145). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-87318-7_7
- Schweiger, W. (2000). Media Credibility — Experience or Image? A Survey on the Credibility of the World Wide Web in Germany in Comparison to other Media. *European Journal of Communication*, *15*(1), 37–59. <https://doi.org/10.1177/0267323100015001002>
- Scott, D. W., & McDonald, C. S. (2005). Utah Newspapers Take Different Approaches to Coverage of First Amendment Controversy. *Newspaper Research Journal*, *26*(2–3), 119–123. <https://doi.org/10.1177/073953290502600210>
- Seckler, M., Heinz, S., Forde, S., Tuch, A. N., & Opwis, K. (2015). Trust and distrust on the web: User experiences and website characteristics. *Computers in Human Behavior*, *45*, 39–50. <https://doi.org/10.1016/j.chb.2014.11.064>
- Self, C. C. (2008). Credibility? In D. W. Stacks & M. B. Salwen (Eds.), *An Integrated Approach to Communication Theory and Research* (Second Edition). Routledge. <http://www.tandfebooks.com/isbn/9780203887011>
- Sendén, M. G., Lindholm, T., & Sikström, S. (2014). Biases in News Media as Reflected by Personal Pronouns in Evaluative Contexts. *Social Psychology*, *45*(2), 103–111. <https://doi.org/10.1027/1864-9335/a000165>
- Sendén, M. G., Sikström, S., & Lindholm, T. (2015). “She” and “He” in News Media Messages: Pronoun Use Reflects Gender Biases in Semantic Contexts. *Sex Roles*, *72*(1–2), 40–49. <https://doi.org/10.1007/s11199-014-0437-x>
- Sexton, C. M. (2014). *Zimmerman and Martin as media’s ammunition: A social constructionist approach to race relations and gun violence* [M.S., Southeastern Louisiana University]. <https://search.proquest.com/docview/1540821177/abstract/1E419A4FB1CD4B2FPQ/1>
- Sherman, J. W., Gawronski, B., & Trope, Y. (2014). *Dual-Process Theories of the Social Mind*. Guilford Publications.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. <https://doi.org/10.1037/0033-295X.84.2.127>
- Shneiderman, B. (2000). Designing Trust into Online Experiences. *Communications of the ACM*, *43*(12), 57–59. <https://doi.org/10.1145/355112.355124>
- Shoemaker, P. J. (1996). *Mediating the Message: Theories of Influences on Mass Media Content*. Longman.
- Shoemaker, P. J., & Reese, S. D. (1996). *Mediating the message: Theories of influences on mass media content* (2nd ed). Longman.
- Shoemaker, P. J., & Vos, T. (2009). *Gatekeeping Theory*. Routledge.
- Simpson, H., & Curtis, J. (2011). U.S. Media Objectivity and Arab Issues: A Content Analysis of Coverage of the Proposed “Ground Zero Mosque.” *The Journal of Undergraduate Research and Creative Activities for the State of North Carolina*, 105–120.
- Siroker, D., & Koomen, P. (2013). *A/B Testing: The Most Powerful Way to Turn Clicks Into Customers*. John Wiley & Sons. https://books.google.com/books?hl=en&lr=&id=VfVvAAAAQBAJ&oi=fnd&pg=PT6&dq=A/B+testing+dan+siroker&ots=T_TYAe7We8&sig=doc3QwD4B-qXlocmmN52VwGUsEQ
- Slater, M. D., & Rouner, D. (1996). How Message Evaluation and Source Attributes May Influence Credibility Assessment and Belief Change. *Journalism & Mass Communication Quarterly*, *73*(4), 974–991. <https://doi.org/10.1177/107769909607300415>

- Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From Protest to Agenda Building: Description Bias in Media Coverage of Protest Events in Washington, D.C. *Social Forces*, 79(4), 1397–1423. <https://doi.org/10.1353/sof.2001.0053>
- Smith, K. B. (1997). When All's Fair: Signs of Parity in Media Coverage of Female Candidates. *Political Communication*, 14(1), 71–82. <https://doi.org/10.1080/105846097199542>
- Smith, N. L. (2014). Consumer Protection in the Marketplace of Ideas: A Proposal to Extend the News Distortion Doctrine to Cable Television News Programs. *Thurgood Marshall Law Review*, 40, 223–272.
- Smith, S. A., & Roden, C. D. (1988). CBS, The New York Times and reconstructed political reality. *Southern Speech Communication Journal*, 53(2), 140–158. <https://doi.org/10.1080/10417948809372719>
- Smucker, M. D., & Jethani, C. P. (2011). The crowd vs. the lab: A comparison of crowd-sourced and university laboratory participant behavior. *Proceedings of the SIGIR 2011 Workshop on Crowdsourcing for Information Retrieval, Beijing*, 194. <http://mpss.csce.uark.edu/smsdb/IPSN/The%20Crowd%20vs.%20the%20Lab:%20A%20Comparison%20of%20Crowd-Sourced.pdf>
- Social Security. (2016, February 20). *Average Wage Index (AWI)*. <https://www.ssa.gov/oact/cola/awidevelop.html>
- Soloski, J. (1989). News reporting and professionalism: Some constraints on the reporting of the news. In Daniel. A. Berkowitz (Ed.), *Media, Culture & Society* (Vol. 11, pp. 207–228).
- Sørensen, J. K. (2013). PSB goes personal: The failure of personalised PSB web pages. *MedieKultur: Journal of Media and Communication Research*, 29(55), 28. <https://doi.org/10.7146/mediekultur.v29i55.7993>
- Soroka, S., & McAdams, S. (2015). News, Politics, and Negativity. *Political Communication*, 32(1), 1–22. <https://doi.org/10.1080/10584609.2014.881942>
- Spillane, B., Hoe, I., Brady, M., Wade, V., & Lawless, S. (2020, April 25). Tabloidization versus Credibility: Short Term Gain for Long Term Pain. *CHI '20: The ACM Conference on Human Factors in Computing Systems*. CHI, Honolulu, Hawai'i, USA. <http://dx.doi.org/10.1145/3313831.3376388>
- Spillane, B., Lawless, S., & Wade, V. (2020, July 19). The Impact of Increasing and Decreasing the Professionalism of News Webpage Aesthetics on the Perception of Bias in News Articles. *Proceedings of the 22nd International Conference On Human-Computer Interaction*. International Conference On Human-Computer Interaction, Copenhagen, Denmark. https://doi.org/10.1007/978-3-030-49059-1_50
- Spillane, B., Lawless, S., & Wade, V. (2018). Increasing and Decreasing Perceived Bias by Distorting the Quality of News Website Design. *Proceedings of the 32nd British HCI Group Annual Conference on People and Computers.*, 13. <https://doi.org/10.14236/ewic/HCI2018.61>
- Spillane, B., Lawless, S., & Wade, V. (2017a). Measuring Bias in News Websites, Towards a Model for Personalization. In M. Bieliková, E. Herder, F. Cena, & M. C. Desmarais (Eds.), *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 387–388). ACM. <https://doi.org/10.1145/3079628.3079647>
- Spillane, B., Lawless, S., & Wade, V. (2017b). Perception of Bias: The Impact of User Characteristics, Website Design and Technical Features. *WI '17: Proceedings of the International Conference on Web Intelligence*, 227–236. <https://doi.org/10.1145/3106426.3106474>
- Spillane, B., & Wade, V. (2020). A Classification of Biases Relating to the Production, Dissemination and Consumption of News. *CHI 2020 Workshop on Detection and Design for Cognitive Biases in People and Computing Systems*, 8. http://critical-media.org/cobi/docs/Spillane_Bias%20Classification%20News.pdf
- Steinbrück, U., Schaumburg, H., Duda, S., & Krüger, T. (2002). A Picture Says More Than a Thousand Words: Photographs As Trust Builders in e-Commerce Websites. *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, 748–749. <https://doi.org/10.1145/506443.506578>

- Stempel III, G. H. (1961). The Prestige Press Covers the 1960 Presidential Campaign. *Journalism Quarterly*, 38(2), 157–163. <https://doi.org/10.1177/107769906103800201>
- Stempel III, G. H. (1965). The Prestige Press in Two Presidential Elections. *Journalism Quarterly*, 42(1), 15–21. <https://doi.org/10.1177/107769906504200102>
- Stempel III, G. H. (1969). The Prestige Press Meets the Third-Party Challenge. *Journalism Quarterly*, 46(4), 699–706. <https://doi.org/10.1177/107769906904600402>
- Stempel III, G. H., & Windhauser, J. W. (Eds.). (1984a). *Newspaper coverage of the 1984 and 1988 campaigns* (Vol. 50). Greenwood Press. <https://cornerstone.lib.mnsu.edu/speaker-gavel/vol50/iss1/2>
- Stempel III, G. H., & Windhauser, J. W. (1984b). The Prestige Press Revisited: Coverage of the 1980 Presidential Campaign. *Journalism Quarterly*, 61(1), 49–55. <https://doi.org/10.1177/107769908406100107>
- Stempel III, G. H., & Windhauser, J. W. (1989). Coverage by the Prestige Press of the 1988 Presidential Campaign. *Journalism Quarterly*, 66(4), 894–919. <https://doi.org/10.1177/107769908906600418>
- Stephens, R. T. (2004). A Framework for the Identification of Electronic Commerce Design Elements That Enable Trust Within the Small Hotel Industry. *Proceedings of the 42Nd Annual Southeast Regional Conference*, 309–314. <https://doi.org/10.1145/986537.986613>
- Stevenson, R. L., Eisinger, R. A., Feinberg, B. M., & Kotok, A. B. (1973). Untwisting the news twisters: A replication of Efron's study. *Journalism Quarterly*, 50(2), 211–219.
- Stevenson, R. L., & Gaddy, G. D. (1984). Bad news' and the Third World. In R. L. Stevenson & D. L. Shaw (Eds.), *Foreign news and the new world information order* (pp. 88–97).
- Stevenson, R. L., & Greene, M. T. (1980). A Reconsideration of Bias in the News. *Journalism Quarterly*, 57(1), 115–121. <https://doi.org/10.1177/107769908005700117>
- Stewart, L. G., Arif, A., & Starbird, K. (2018). Examining Trolls and Polarization with a Retweet Network. *Proceedings of ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web*, 6.
- Stocking, S. H., & Gross, P. H. (1989). *How Do Journalists Think? A Proposal for the Study of Cognitive Bias in Newsmaking*. ERIC Clearinghouse on Reading and Communication Skills, Smith Research Center, Indiana University.
- Stoneman, P. (2008). *This Thing Called Trust: Civic Society in Britain*. Springer.
- Stoodley, B. H. (1960). Bias in Reporting The FCC Investigation. *Public Opinion Quarterly*, 24(1), 92–98. <https://doi.org/10.1086/266933>
- Sundar, S. S. (1999). Exploring Receivers' Criteria for Perception of Print and Online News. *Journalism & Mass Communication Quarterly*, 76(2), 373–386. <https://doi.org/10.1177/107769909907600213>
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). http://www.marketingsociale.net/download/modello_MAIN.pdf
- Sundar, S. S., Knobloch-Westerwick, S., & Hastall, M. R. (2007). News cues: Information scent and cognitive heuristics. *Journal of The American Society for Information Science and Technology*, 58(3), 366–378. <https://doi.org/10.1002/asi.20511>
- Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, 51(1), 52–72. <https://doi.org/10.1111/j.1460-2466.2001.tb02872.x>
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2), 143–148. [https://doi.org/10.1016/0001-6918\(81\)90005-6](https://doi.org/10.1016/0001-6918(81)90005-6)
- Swerts, M., & Krahmer, E. (2010). Visual prosody of newsreaders: Effects of information structure, emotional content and intended audience on facial expressions. *Journal of Phonetics*, 38(2), 197–206. <https://doi.org/10.1016/j.wocn.2009.10.002>

- Sztompka, P. (1999). *Trust: A Sociological Theory*. Cambridge University Press.
- Tan, Y.-H., & Thoen, W. (2000). Toward a Generic Model of Trust for Electronic Commerce. *International Journal of Electronic Commerce*, 5(2), 61–74. <https://doi.org/10.1080/10864415.2000.11044201>
- Tannenbaum, P. H. (1953). The Effect of Headlines on the Interpretation of News Stories. *Journalism Bulletin*, 30(2), 189–197. <https://doi.org/10.1177/107769905303000206>
- Tanner, A. H. (2004). Agenda Building, Source Selection, and Health News at Local Television Stations: A Nationwide Survey of Local Television Health Reporters. *Science Communication*, 25(4), 350–363. <https://doi.org/10.1177/1075547004265127>
- Tarafdar, M., & Zhang, J. (2005). Analysis of critical website characteristics: A cross-category study of successful websites. *Journal of Computer Information Systems*, 46(2), 14–24.
- The Irish Times View. (2018, March 6). In a spin: Leo Varadkar's Strategic Communications Unit. *The Irish Times*. <https://www.irishtimes.com/opinion/editorial/in-a-spin-leo-varadkar-s-strategic-communications-unit-1.3415732>
- Towner, T. L. (2017). The Role of Visual Content on Social Media in the 2016 Presidential Campaign. In D. Schill & J. A. Hendricks (Eds.), *The Presidency and Social Media: Discourse, Disruption, and Digital Democracy in the 2016 Presidential Election*. Routledge.
- Turner, J. (2007). The Messenger Overwhelming the Message: Ideological Cues and Perceptions of Bias in Television News. *Political Behavior*, 29(4), 441–464. <https://doi.org/10.1007/s11109-007-9031-z>
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Tyler Eastman, S. (2001). Biased voices of sports: Racial and gender stereotyping in college basketball announcing. *Howard Journal of Communication*, 12(4), 183–201.
- Updegraff, J. A., Sherman, D. K., Luyster, F. S., & Mann, T. L. (2007). The effects of message quality and congruency on perceptions of tailored health communications. *Journal of Experimental Social Psychology*, 43(2), 249–257.
- Jacobellis v. Ohio, 378 U.S. 184 (1964), (U.S. Supreme Court June 22, 1964). <https://supreme.justia.com/cases/federal/us/378/184/case.html>
- Van Belle, D. A. (2000). New York Times and Network TV News Coverage of Foreign Disasters: The Significance of the Insignificant Variables. *Journalism & Mass Communication Quarterly*, 77(1), 50–70. <https://doi.org/10.1177/107769900007700105>
- Van Dijk, T. A. (2000). New(s) Racism: A discourse analytical approach. *Ethnic Minorities and the Media*, 37, 33–49.
- Van Hout, T., & Jacobs, G. (2008). News production theory and practice: Fieldwork notes on power, interaction and agency. *Pragmatics*, 18(1), 59–85. <https://doi.org/10.1075/prag.18.1.04hou>
- VandeHei, J., & Allen, M. (2012, May 31). To GOP, blatant bias in vetting. *POLITICO*. <https://www.politico.com/news/stories/0512/76898.html>
- Vergeer, M., Lubbers, M., & Scheepers, P. (2000). Exposure to Newspapers and Attitudes toward Ethnic Minorities: A Longitudinal Analysis. *Howard Journal of Communications*, 11(2), 127–143. <https://doi.org/10.1080/106461700246661>
- Waldman, P., & Devitt, J. (1998). Newspaper photographs and the 1996 presidential election: The question of bias. *Journalism & Mass Communication Quarterly*, 75(2), 302–311.
- Walther, J. B., Wang, Z., & Loh, T. (2004). The Effect of Top-Level Domains and Advertisements on Health Web Site Credibility. *Journal of Medical Internet Research*, 6(3). <https://doi.org/10.2196/jmir.6.3.e24>
- Wang, X. (2003). *Media ownership and objectivity* [Masters, Louisiana State University - The Manship School of Mass Communication]. https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=4401&context=gradschool_theses

- Wang, Y. D., & Emurian, H. H. (2005). An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior, 21*(1), 105–125. <https://doi.org/10.1016/j.chb.2003.11.008>
- Wang, Z., Walther, J. B., Pingree, S., & Hawkins, R. P. (2008). Health Information, Credibility, Homophily, and Influence via the Internet: Web Sites Versus Discussion Groups. *Health Communication, 23*(4), 358–368. <https://doi.org/10.1080/10410230802229738>
- Warhurst, S., McCabe, P., & Madill, C. (2013). What Makes a Good Voice for Radio: Perceptions of Radio Employers and Educators. *Journal of Voice, 27*(2), 217–224. <https://doi.org/10.1016/j.jvoice.2012.08.010>
- Watanabe, K. (2017a). *Measuring bias in international news: A large-scale analysis of news agency coverage of the Ukraine crisis* [PhD Thesis]. The London School of Economics and Political Science (LSE).
- Watanabe, K. (2017b). Measuring news bias: Russia’s official news agency ITAR-TASS’ coverage of the Ukraine crisis. *European Journal of Communication, 32*(3), 224–241. <https://doi.org/10.1177/0267323117695735>
- Wathen, C. N., & Burkell, J. (2002). Believe It or Not: Factors Influencing Credibility on the Web. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 53*(2), 134–144. http://www.communicationcache.com/uploads/1/0/8/8/10887248/believe_it_or_not_-_factors_influencing_credibility_on_the_web.pdf
- Weatherly, J. N., Petros, T. V., Christopherson, K. M., & Haugen, E. N. (2007). Perceptions of Political Bias in the Headlines of Two Major News Organizations. *Harvard International Journal of Press/Politics, 12*(2), 91–104. <https://doi.org/10.1177/1081180X07299804>
- Weaver, J. B., Huck, I., & Brosius, H.-B. (2009). Biasing public opinion: Computerized continuous response measurement displays impact viewers’ perceptions of media messages. *Computers in Human Behavior, 25*(1), 50–55. <https://doi.org/10.1016/j.chb.2008.06.004>
- Weaver, P. H. (1972). Is Television News Biased? *The Public Interest, 26*, 57–74.
- Weibel, D., Wissmath, B., & Groner, R. (2008). How Gender and Age Affect Newscasters’ Credibility—An Investigation in Switzerland. *Journal of Broadcasting & Electronic Media, 52*(3), 466–484. <https://doi.org/10.1080/08838150802205801>
- Weisbuch, M., Pauker, K., & Ambady, N. (2009). The Subtle Transmission of Race Bias via Televised Nonverbal Behavior. *Science, 326*(5960), 1711–1714. <https://doi.org/10.1126/science.1178358>
- Westerwick, A. (2013). Effects of Sponsorship, Web Site Design, and Google Ranking on the Credibility of Online Information. *Journal of Computer-Mediated Communication, 18*(2), 80–97. <https://doi.org/10.1111/jcc4.12006>
- White, D. M. (1950). The “Gate Keeper”: A Case Study in the Selection of News. *Journalism Bulletin, 27*(4), 383–390. <https://doi.org/10.1177/107769905002700403>
- Whitehead, J. L. (1968). Factors of source credibility. *Quarterly Journal of Speech, 54*(1), 59–63. <https://doi.org/10.1080/00335636809382870>
- Whitney, D. C., Fritzier, M., Jones, S., Mazzarella, S., & Rakow, L. (1989). Geographic and source biases in network television news 1982-1984. *Journal of Broadcasting & Electronic Media, 33*(2), 159–174. <https://doi.org/10.1080/08838158909364070>
- Wilhoit, G. C., & Auh, T. S. (1974). Newspaper Endorsement and Coverage of Public Opinion Polls in 1970. *Journalism Quarterly, 51*(4), 654–658. <https://doi.org/10.1177/107769907405100411>
- Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*(8), 594.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*(1). <https://doi.org/10.1038/sdata.2016.18>

- Williams, A. (1975). Unbiased Study of Television News Bias. *Journal of Communication*, 25(4), 190–199. <https://doi.org/10.1111/j.1460-2466.1975.tb00656.x>
- Wilson, P. (1983). *Second-Hand Knowledge: An Inquiry Into Cognitive Authority*. Greenwood Press.
- Wisker, G. (2007). *The Postgraduate Research Handbook: Succeed with Your MA, MPhil, EdD and PhD*. Palgrave Macmillan.
- Wolfenbarger, M., & Gilly, M. (2002). . comQ: Dimensionalizing, measuring, and predicting quality of the e-tail experience. *Marketing Science Institute Report*, 02–100.
- Wolfenbarger, M., & Gilly, M. C. (2003). eTailQ: Dimensionalizing, measuring and predicting etail quality. *Journal of Retailing*, 79(3), 183–198. [https://doi.org/10.1016/S0022-4359\(03\)00034-4](https://doi.org/10.1016/S0022-4359(03)00034-4)
- Wolton, S. (2017). Are Biased Media Bad for Democracy? *American Journal of Political Science*. <https://mpr.aub.uni-muenchen.de/84837/>
- Wortham, S., & Locher, M. (2009). Voicing on the news: An analytic technique for studying media bias. *Text - Interdisciplinary Journal for the Study of Discourse*, 16(4), 557–586. <https://doi.org/10.1515/text.1.1996.16.4.557>
- Yalch, R. F., & Elmore-Yalch, R. (1984). The Effect of Numbers on the Route to Persuasion. *Journal of Consumer Research*, 11(1), 522–527.
- Yamamoto, Y., & Tanaka, K. (2011). Enhancing Credibility Judgment of Web Search Results. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1235–1244. <https://doi.org/10.1145/1978942.1979126>
- Yao-Hua Tan, W. T. (2000). Toward a Generic Model of Trust for Electronic Commerce. *International Journal of Electronic Commerce*, 5(2), 61–74. <https://doi.org/10.1080/10864415.2000.11044201>
- Young, A. L., Komlodi, A., Rózsa, G., & Chu, P. (2016). Evaluating the credibility of English web sources as a foreign-language searcher. *Proceedings of the Association for Information Science and Technology*, 53(1), 1–9. <https://doi.org/10.1002/pr.a2.2016.14505301042>
- Zaller, J. (1999). *A Theory of Media Politics: How the Interests of Politicians, Journalists, and Citizens Shape the News*. <https://polisci.ucla.edu/people/john-zaller>. https://polisci.ucla.edu/sites/default/files/u244/theory_of_media_politics_2017.pdf
- Zaller, J. R. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.
- Zaller, J. R. (1999). A theory of media politics: How the interests of politicians, journalists, and citizens shapes the news. *University of California, Los Angeles. Typescript*.
- Zeldes, G. A., Fico, F., Carpenter, S., & Diddi, A. (2008). Partisan Balance and Bias in Network Coverage of the 2000 and 2004 Presidential Elections. *Journal of Broadcasting & Electronic Media*, 52(4), 563–580. <https://doi.org/10.1080/08838150802437354>
- Zelizer, B., Park, D., & Gudelunas, D. (2002). How bias shapes the news: Challenging The New York Times' status as a newspaper of record on the Middle East. *Journalism*, 3(3), 283–307.
- Zhang, P., Dran, G. von, Blake, P., & Pipithsukunt, V. (2000). A comparison of the most important website features in different domains: An empirical study of user perceptions. *Americas Conference on Information Systems - AMCIS 2000 Proceedings*, 41.
- Zhang, P., & von Dran, G. M. (2000). Satisfiers and dissatisfiers: A two-factor model for website design and evaluation. *Journal of the American Society for Information Science*, 51(14), 1253–1268. [https://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-AS11039>3.0.CO;2-O](https://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-AS11039>3.0.CO;2-O)
- Zimmerman, J. (2013). *Media Bias through Facial Expressions on Local Las Vegas Television News*. University of Nevada.

9 Appendices

9.1 Appendix A – Chapter One: Bias in the Research

Five main sources of possible participant bias have been identified which may impact on this research. 1) All the participants surveyed were based in the US and their first language was English. They are also predominately, young, well-educated and affluent. They also all had access to the WWW and the majority report using it often, both on their phone and other devices. 2) All the participants were paid to undertake the experiments and sourced through an online marketplace Prolific Academic. Having knowledge of, or access to such a marketplace is also indicative of a technically literate, educated sample of participants. 3) It is likely that the majority of the participants would also have more than a passing interest in the news and/or bias, and how it affects news consumption. This was shown by the large uptake in each experiment. It is likely that such participants are more keenly aware and opinionated as to the type and quality of news they consume. 4) It is also possible that some of the participants had strong feelings towards some of the websites or the topics in the articles in the experiments despite every precaution taken against it. 5) There is also a slight concern that participants did not conduct themselves as supposed. The most obvious possible manifestation of this is in the bias rating tasks. In each, participants were required to rate the perceived bias in each distorted webpage/article combination. The instructions to do so were very clear and explicit and reiterated multiple times. However, it is possible, though unlikely, that participants ignored, consciously or unconsciously, the instructions and instead rated the design or some other aspect of the distorted webpage/article combinations. Essentially if you give a participant a button, they will press it. To combat this, several validation methods and attention tasks were implemented in each experiment.

Two main forms of researcher bias have been identified despite taking all precautions against them. 1) Personal biases which may have manifested itself in the choice of websites, the topics covered in the news articles and the analysis of results. In each case the details are presented in chapters Four, Five and Six, so that the reader may make their own judgements. 2) The distorted webpage/article combinations were created by the researcher based on their suppositions which in turn were based on their reading of the literature in the domain. It is possible that these are unrealistic or distorted by the researchers own opinions of online news access.

9.2 Appendix B – Chapter Two: Definitions of Bias

9.2.1.1.1 Definitions of Media Bias

Media is the overarching domain which includes all television, radio, traditional print, and WWW communication. While most think of bias in terms of news, it can also be present in other areas such as economic reporting, photographs, advertising, election coverage, documentaries, and sports commentary etc. For his study of racial bias in network television coverage of American football games, Rada grounded his study in the earlier work of Rainville and McCormick's study into racial prejudice in professional American football commentators' speech. Rada's study used a Biased Coverage Index to characterise descriptive comments of each player by announcers to ascertain if racial bias was present. While neither Rada nor Rainville and McCormick provide an explicit lexical definition of the construct, Rada's description of bias used in the previous study forms an early quasi-definition and was among the first to highlight the positive / negative nature of certain aspects of bias. *"For this study, the definition of bias is grounded in Rainville and McCormick's (1977) research. Rainville and McCormick concluded that announcers gave biased coverage by presenting a negative image of African American players while presenting a positive image of White players."* (Rada, 1996), citing (Rainville & McCormick, 1977). The study found that commentators were more likely to emphasise the athleticism of African American players while Caucasian players were more likely to have their cognitive ability spoken of favourably.

One of the most influential scholars in mass communication, McQuail defined bias as: *"a consistent tendency to depart from the straight path of objective truth by deviating either to left or right"* (McQuail, 1992). His book relates specifically to concerns about the nature and quality of news media and its serving of the public interest. It highlights different means of measuring media performance, including perceived bias. While this definition is highly regarded and oft cited in the core literature, it has not been used in any empirical study encountered in this review. This is likely due to its limitations in that it specifically relates to 'objective truth', thus ignoring other forms of bias such as coverage and selection bias and due to the use of the terms 'left' and 'right' which means it could only be used in studies relating to conservative and liberal coverage. McQuail maintains that bias has four different forms: partisanship, unwitting, ideology and propaganda (McQuail, 1992).

In their analysis of the market for news media, the economists Mullainathan and Shleifer do not offer a specific lexical definition of media bias. They define bias in terms of accuracy rather than untruth, where a journalist fails to report a fact rather than chooses to report a falsehood (Mullainathan & Shleifer, 2005). They argue that this form of media bias is undertaken by media organisations to suit the preferences of their audiences. They maintain that media organisations are engaged in 'slanting', where journalists may choose to omit relevant facts. Slanting was first defined by Hayakawa as: *"the process of selecting details favourable or unfavourable to the subject being described"* (Hayakawa, 1939, 1949). Mullainathan and Shleifer's model demonstrates that on subjects where the readers have a common position, the media slants

towards the consensus. However, this may not be the same as an accurate depiction of the facts or situation. This is undertaken in the knowledge that news consumers form higher opinions of news and news organisations that produce news in accordance with their beliefs. More importantly they also show that on topics where readers beliefs diverge, newspapers segment the market by slanting towards extreme or fringe positions. They also note that by consuming media from different publications, it is possible to get an unbiased perspective in the aggregate.

Groseclose and Milyo offer an all-encompassing definition based on ideological slant with: *“we define bias as an ideological slant that may take a number of forms: Democratic or Republican partisanship; liberal or conservative positions on public-policy issues; or broader assumptions about, say, business corporations or the causes of social, economic, and foreign-policy problems”* (Groseclose & Milyo, 2005b). This definition recognises the limiting nature of focusing on accuracy, honesty, trust, or on positive or negative coverage etc., and instead hints that bias may be all or part of these things.

This is reflected in Groseclose and Milyo’s subsequent innovative content analysis study published later the same year (Groseclose & Milyo, 2005a). It recorded the number of times certain media outlets cite various think tanks and policy groups and then compared that to the number of times conservative and liberal members of the US congress cited the same groups. This was based on the supposition that liberal politicians cite liberal think tanks and policy groups and conservative politicians cite conservative think tanks and policy groups. The results showed a strong liberal bias in all the major media outlets except for Fox News and the Washington Times. In contrast to McQuail, Groseclose and Milo specifically reject a definition of bias based on truth, accuracy, or honesty, stating that their definition: *“has nothing to do with the honesty or accuracy of a news outlet. Instead our notion is more like a taste or preference.”* Understanding how each participant may have a different understanding of the term they go on to say: *“In contrast, other writers, at least at times, do define bias as a matter of accuracy or honesty. We emphasize that our differences with such writers are ones of semantics, not substance. If, say, a reader insists that bias should refer to accuracy or honesty, then we urge him or her simply to substitute another word wherever we write ‘bias.’ Perhaps ‘slant’ is a good alternative.”* (emphasis on ‘do’ in original text) (Groseclose & Milyo, 2005a). Their rejection of a definition based on accuracy or honesty is based on their claim that journalists rarely make dishonest statements, they go on to state: *“we argue that our notion of bias is meaningful and relevant, and perhaps more meaningful and relevant than the alternative notion”*. They further maintain: *“Instead, for every sin of commission... ..we believe that there are hundreds, and maybe thousands, of sins of omission”*.

The economists Gentzkow and Shapiro, in another of the most highly cited works in the domain, provide three written accounts from Fox News, The New York Times, and Al Jazeera of the same battle between US troops and insurgents in the city of Samarra in northern Iraq on December 2nd, 2003. They demonstrate

that despite all three accounts being based on the same set of underlying facts, the message transmitted by the content to the reader can be vastly different. They highlight these inconsistencies between accounts when explaining what they mean by bias: *“All the accounts are based on the same set of underlying facts. Yet by selective omission, choice of words, and varying credibility ascribed to the primary source, each conveys a radically different impression of what actually happened. The choice to slant information in this way is what we will mean in this paper by media bias.”* (Gentzkow & Shapiro, 2006). This definition is noteworthy as it also highlights the role of credibility. They built a model based on the fact that media firms slant their news towards the prior opinions of their readers in order to increase perceived credibility. Their model reveals this bias, despite its damage to the news organisations reputation. Gentzkow and Shapiro’s work had similar findings to two other economists Mullainathan and Shleifer, who found that bias is more prevalent in sectors of the news where there is a larger diversity of opinion than when there is common agreement, further proving that media organisation bias news to suit their consumers’ opinions (Mullainathan & Shleifer, 2005). In subsequent work Gentzkow, Shapiro and Stone review the theoretical literature on the market influences on bias and introduce a model on the two determinates of bias, supply side bias or demand side bias (Gentzkow et al., 2015).

In one of many tests for partisan bias between the major news networks in the US, Groeling offers a succinct yet encompassing definition unbounded by the domain. *“In this analysis, I define media bias as a portrayal of reality that is significantly and systematically (not randomly) distorted.”* (parentheses in original text) (Groeling, 2013a). His work was designed to overcome perceived deficiencies in some of the previous studies in the domain, most notably their inability to accurately test for selection bias. It tackles this issue by highlighting the reporting of presidential approval polls across the major news networks. Unlike prior works in the domain, Groeling was forearmed with the knowledge of all the major polls conducted. This enabled him to identify when particular news agencies chose to highlight or ignore polls which they perceived as being good or bad for the president they supported. This study demonstrated that conservative and liberal leaning news agencies tended to highlight, bury or even ignore polls which did not provide favourable coverage to the conservative or liberal presidents they supported. It also enabled him to measure how much attention they gave to polls they chose to highlight. This work should be studied in unison with an earlier study by Wilhoit and Auh who found: *“a significant relationship between newspaper endorsement of candidates and favorable opinion poll coverage about them.”* (Wilhoit & Auh, 1974).

Hoffman citing Lee maintains that: *“Media bias is often defined as the opposite of objectivity, in that the content can be measured as being favorable toward some side of an issue.”* (Hoffman, 2013; T.-T. Lee, 2008). Her study focused on so called ‘new news formats’ such as political satire shows or cable news talk shows. Participants were randomly assigned to watch an interview with a 2012 US presidential candidate on one of these formats. Using a combination of questions and a scale to measure bias, the study found significant differences amongst the participants in perceived bias towards the candidate.

9.2.1.1.2 *Definitions of Political Bias*

This literature review discovered only two definitions specific to political bias that were not specifically or lexically related to news. The first was put forward by Russo in his formative study on television news bias in the coverage of the war in Vietnam from 1969 to 1970. *“The concept of ‘bias’ is key to the study. It was defined as that quality of statements of opinion or of actual or supposed fact that would influence one to support or oppose a President or his policy.”* (Russo, 1971). This definition is unique at the time for its attempt to ignore whether a statement was true or without omission, but whether or not it could influence ones’ opinion. He goes on to state: *“The veracity of a given statement is thus independent of its ‘bias’ as defined above, and a neutral rating on the scale measuring bias... ..is not ipso facto desirable, ‘fair,’ or objective reporting. Bias defined in this manner enables one to see differences between the networks and to gauge the over-all impact of the network. From this, he can decide whether or not he considers this coverage ‘fair.’”* (Russo, 1971). The study found no discernible pattern of bias against the Nixon administrations’ policies in Vietnam by the two US television networks tested, NBC or CBS.

Waldman and Devitt’s defined of bias as: *“any systematic slant favoring one candidate or ideology over another”*. (Waldman & Devitt, 1998). Their study of pictorial treatment of Bill Clinton and Bob Dole during the 1996 US presidential campaign found that Clinton received slightly better treatment overall and that the favourability of pictures towards both candidates rose and fell depending on the polls. The authors argue against a liberal media bias and instead argue in favour of a frontrunner bias. This is just one of many studies into bias in election photographs used in print publications, others include; (Barrett & Barrington, 2005; Greenwood, 2005; Moriarty & Garramone, 1986; Moriarty & Popovich, 1991).

A quasi definition of political bias by Diddi et al. was provided in their contribution on partisan bias in television coverage of three US presidential elections. Their consideration of news bias is limited to political bias and in many ways lacks the careful crafting of many of the other definitions presented here: *“this study considered news bias a causal phenomenon within news outlets that produces unequal treatment of candidates in a political race.”* (Diddi et al., 2014).

9.2.1.1.3 *Definitions of News Bias*

The definition of bias used in Edith Efron’s seminal work was based on the Fairness Doctrine. This was introduced in 1949 by the Federal Communications Commission (FCC) as a modification of its earlier 1941 ruling known as the Mayflower Decision. The Fairness Doctrine had two basic elements. It required broadcasters to dedicate some of their time to controversial matters of public interest, and that they had to

air multiple different viewpoints on those matters. This was revoked in 1987⁴¹. The Fairness Doctrine essentially defines bias as the opposite of fairness. Efron maintains that the Fairness Doctrine holds that:

- *“The networks are required to select and broadcast contrasting and conflicting views on the major political issues - regardless of their truth or falsity.”*
- *“This selective process is to be ‘nonpartisan’ and ‘non-one-sided,’ i.e., favoring neither side.”*
- *“And the selected opinion must be presented in an ‘equal’ and ‘equally forceful’ manner.”*

“To do this, says the FCC, is fairness. To fail to do this, says the FCC, is bias.” Efron later states: *“The FCC definition... ..is the one formally used by this study.”* (Efron, 1971, p. 6).

However, it should be noted, and as Williams points out, a major failing of the work is that Efron herself did not put forward a specific lexical definition of bias (Williams, 1975). This is also pointed out by Doll and Bradley who maintain that in Efron’s understanding, bias has three dimensions; selectivity; the presence of opinion; and what she believes to be the two main points contained in the FCC doctrine, reasonable temporal equality on issues, and balance in opinion (Doll & Bradley, 1974). It should also be noted that Efron’s study is not without its detractors, including Doll and Bradley and Stevenson et al. In their rerunning of her experiment Stevenson et al. highlight three basic methodological shortcomings (Stevenson et al., 1973), while Doll and Bradley note the: *“lack of controlled objectivity”*, *“questionable use of clearly labeled editorial opinion”*, and *“a confusion in methodology”*. It should be noted however, that as one of the earliest works in the domain and considering her training and career as a journalist and not as a researcher, her study was a catalyst for further research in the domain⁴².

US presidential campaigns have proved a mainstay for studies into news bias. In his study of network bias in the 1972 presidential campaign Lowry uses an earlier definition put forward in a study by Klein and Maccoby into Newspaper objectivity in the 1952 presidential campaign. *“‘Bias’ is here defined as the existence of a differential, larger than could be expected by chance alone, between the proportional front-page coverage allotted the two candidates by the two sets of papers. That is, bias is defined as the difference between proportions, or means”* (Klein & Maccoby, 1954; D. T. Lowry, 1973). In their highly cited meta-

⁴¹ The Fairness Doctrine was opposed by many leading conservative and libertarians who believed it was an attack on the First Amendment. To read more on how and why it was revoked see “The Federal Communications Commission, the Broadcast Industry, and the Fairness Doctrine: 1981-1987” (D. J. Jung, 1996).

⁴² Weaver is one of the few to defend Efron’s contribution. He states his disappointment with the community that the book was not received and treated fairly, or recognised as a serious independent effort, though not without its drawbacks. *“One could have hoped, therefore, that Edith Efron’s The News Twisters [] would have been recognized for what it is -- a serious independent effort to measure what she takes to be the ‘fairness’ of network Television news, and thereby to shed the light of quantitative evidence on one important aspect of the debate. Alas, it generally has been recognized as no such thing, and both its supporters and detractors have quickly assimilated it to their own polemical purposes. This is a pity, for the book deserves more thoughtful attention than it has generally received so far -- not so much because it is brilliant, or scholarly, or especially penetrating (it is none of these), but rather because it presents for the first time an extensive and tolerably reliable body of evidence on the content of Television coverage of the presidential campaign in 1968. Properly interpreted, this evidence reveals a great deal about how Television news is biased and thus about the direction of its influence on public”* (P. H. Weaver, 1972).

analysis of media bias in presidential elections, D'Alessio and Allen categorise three types of bias as; gatekeeping, coverage, and statement bias (D. D'Alessio & Allen, 2000). The studies by Lowry, and Klein and Maccoby fall into the category of coverage bias, literally the amount of time, space, or column inches that is dedicated to one candidate, party, or ideology. Coverage bias is a serious issue and has been studied in some format in every single US presidential election since at least 1952 (Klein & Maccoby, 1954; Kobre, 1953). Though several of the studies found examples of coverage bias, Stevenson and Greene found little evidence of what they termed: "*large scale, systematic favoring of one political party or candidate over the other in the last two or three decades.*" (Stevenson & Greene, 1980). Many studies on coverage bias were undertaken by Stempel and Windhauser and relate directly to column inches (Stempel III, 1961, 1965, 1969; Stempel III & Windhauser, 1984a, 1984b, 1989). D'Alessio and Allen point out that there are also studies that have also been undertaken on the size and favourability of photos (Barrett & Barrington, 2005; Batlin, 1954; Greenwood, 2005; Kenney & Simpson, 1993; Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Waldman & Devitt, 1998), and on headline bias (Weatherly et al., 2007). As Cohen stated so eloquently: "*the media may not be successful much of the time in telling people what to think, but is stunningly successful in telling its readers what to think about*" (emphasis on 'about' in original text) (B. C. Cohen, 1963, p. 13). Or as Brodie et al. citing Mrogers and Wdearing state: "*most scholars would agree that the media do, indeed, have an indirect effect on the public agenda; that is, people get a sense from the media about which agenda items are more important than the others: 'not what to think, but what to think about.'*" (Brodie et al., 1998; Mrogers & Wdearing, 1988).

In their study of network television coverage of the 1972 presidential campaign, Doll and Bradley put forward three cautions before they offer forth a definition of bias. First, is the abundance of synonyms for the term. Second, is the fact that bias itself has become a negatively loaded term, which is: "*used as automatic invective by people who dislike the networks on political grounds.*" Third, the fundamentally almost ineffable nature of the construct causing them to state: "*the term is abstract; therefore, it tends to mean, like 'democracy,' what people want it to mean. At worst, bias becomes 'what I say it is.'*" (Doll & Bradley, 1974). Consequently, they attempt to avoid these difficulties by rooting their definition of bias in objectivity: "*In order to avoid these difficulties, it was determined that bias be negatively defined as the absence of objectivity, and that the term objectivity be operationally defined in a series of questions.*" (emphasis on 'absence of objectivity' in original text) (Doll & Bradley, 1974).

Following along the same lines as the earlier books on the subject, Hofstetter also does not put forward a specific lexical definition. The closest he comes to defining bias is with: "*Bias is a helpful term if we are careful of how we use it. In both lay usage and scientific writing 'bias' encompasses many meanings. The layman uses 'bias,' along with such terms as 'slanted reporting' and 'coloring the facts,' to put a wide range of practices into one package. These practices may include outright falsehood at one extreme. 'Bias' may also include analyses and evaluations that the newsman will defend as legitimate expressions of*

personal judgment. The scientist finds the word helpful for directing the reader's attention to a precise area of inquiry. This report always uses 'bias' to denote partiality in news programming" (Hofstetter, 1976). His cautious approach and differentiation in understandings of the concept underscores the difficulty in defining the term and serves as a caveat to his subsequent definition. It is also noteworthy for its use of the term 'partiality', without highlighting any systematic approach or pattern of partiality which subsequently became the norm in other definitions.

Stevenson and Green maintain that the standard definition of news bias, and one that has been adopted in most serious news studies up to that point was: "*bias is the systematic differential treatment of one candidate, one party, one side of an issue over an extended period of time. Bias is the failure to treat all voices in the marketplace of ideas equally*" (Stevenson & Greene, 1980). However, they argued against this definition and suggest a "*reconsideration*" of the construct of bias. Mantler and Whitman also highlight that Stevenson and Green did not agree with the standard definition. "*They argued this definition is flawed, and newspapers, in their attempt to be fair, may force equal coverage of two candidates and distort the real differences in the candidates and their campaigns and suggested using specific cases in discussing bias; contending that one problem with past studies has been the examinations of the papers were over relatively long periods of time, enabling papers to balance their overall coverage, even if specific issues might show tremendous inequalities.*" (Mantler & Whiteman, 1995).

Taking an alternative approach to the majority in the domain, Kline, citing Philo et al. states: "*More recently under the pressure of refinements of theory, bias has come to be more broadly defined as an implicit cultural perspective, or 'world view' that is related in the news.*" (Kline, 1981; Philo et al., 1977). In this assessment of the national perspective and comparison of national news broadcasts, Kline centres his definition on the consumer of news, or society in general, rather than on the producers or purveyors of news. This study is interesting in its focus on balance, inference, and facts in a story rather than the much easier to measure column inches, or headline or photograph favourability. However, this contribution to the literature has received little attention.

A succinct definition: "*selective inaccuracy in news coverage*" was put forward by Gaddy and Tanjong in their study of geographical bias in western medias' coverage of earthquakes (Gaddy & Tanjong, 1986). The study showed that restraint must be shown before assigning bias to unequal coverage by western media. They base their definition on the earlier work of Stevenson, who reran Efron's study, and who Gaddy later produced a book chapter with (Efron, 1971; Stevenson et al., 1973; Stevenson & Gaddy, 1984). It must be noted however that this definition has received little attention from other studies into bias, or from the arch domain of credibility. This is likely due to its focus on 'inaccuracy', whereas most other researchers have instead focused on, omission, imbalance, slant, and favourability.

In one of several studies into bias in the facial expressions of television news anchors, Miller et al. specifically defines bias in terms of the positive or negative effects of facial expressions: “*From previous studies we adopt the conceptual definitions of bias as a systematic slant (Waldman & Devitt, 1998), differential amounts of negative, positive, and neutral content (Moriarty & Popovich, 1991), specifically the positivity and negativity of facial expressions (Friedman et al., 1980; Mullen et al., 1986). We adapt these definitions for this study to define bias as differential positive and negative nonverbal behaviors of news anchors. Nonverbal behavior is defined as facial expressions (Englis, 1994) for this study.*” (Miller et al., 2007). This is similar to Russo’s and Stevenson and Green’s aforementioned definitions (Russo, 1971; Stevenson & Greene, 1980). This longitudinal study found that some news anchors exhibited significantly more positive facial expressions when they mentioned one presidential candidate. Interestingly the study points out that viewers who perceive bias, even when none has been found: “*may be picking up on the more subtle, nonverbal elements of media coverage that are captured only in visual studies.*” (Miller et al., 2007). This could include the visual elements or graphics accompanying a news report or additional characteristics of the news caster such as their posture or hand gestures.

In what is likely the most cited definition put forward in a journal article in this State of the Art, despite its relatively recent publication, Entman delineates and differentiates between the separate but connected terms of slant and bias by offering definitions for both. This has long been a source of confusion in the domain with many basing their definitions of bias off slant without defining that term and demarcating the difference. As Entman clearly states: “*To help avoid the terminological confusion discussed previously, I propose to distinguish bias from news slant. Slant characterises individual news reports and editorials in which the framing favors one side over the other in a current or potential dispute.*” (emphasis in original text) (Entman, 2007, p. 165). This definition defines slant as a single, rare or occasional occurrence in favour of one candidate or ideology. Given how difficult if not impossible it is to make every news report completely fair and non-partisan, it could be argued that each has some degree of slant. In contrast, bias is a systematic or consistent pattern of slant which is unlikely to occur by accident. “*This brings us to the proposed definition of content bias: “consistent patterns in the framing of mediated communication that promote the influence of one side in conflicts over the use of government power”*⁴³. By this definition, to reveal media content biases, we must show patterns of slant that regularly prime audiences, consciously or unconsciously, to support the interests of particular holders or seekers of political power.” (emphasis in original text) (Entman, 2007, p. 166).

⁴³ Entman’s research on bias focuses on the media’s agenda setting and framing of problems: “*worthy of public and government attention*” (Entman, 2007, p. 164). He maintains that each form of media bias is a strategic framing contest where the media attempts to influence day to day contests to control government power. He goes on to say: “*Some researchers in the critical studies tradition might conclude that the media meet the suggested standards for bias at a more fundamental level: consistent framing in favor of capitalism, patriarchy, heterosexism, individualism, consumerism, and White privilege, among other deeply entrenched values that certainly help allocate power in American society (Budd, Craig, & Steinman, 1999). However, this research is more concerned with media interventions in the day-to-day contests to control government power within the snug ideological confines of mainstream American politics.*” (Entman, 2007, p. 166).

In their study of the impact of government advertising on coverage of corruption scandals in the four main newspapers in Argentina from 1998 to 2007, Di Tella and Franceschelli do not put forward a specific lexical definition of bias but rather more of a technical method based on the earlier work of Groseclose and Milyo, and Gentzkow and Shapiro. *“In our case, we calculate an average reporting of corruption... ..and observe if newspaper reporting is different than this average when government advertising is relatively high. Thus, if all papers are equally biased, we do not detect it with our tests.”* (Di Tella & Franceschelli, 2011). Their research found some striking results. The main finding being: *“a one standard deviation increase in monthly government advertising is associated with a reduction in the coverage of the government’s corruption scandals of 0.23 of a front page per month, or 18 percent of a standard deviation in coverage.”* Similar results have been found in other media markets such as Italy, where a significant and large correlation was found between positive coverage of listed companies and the advertising expenditure of that company in each newspaper (Gambaro & Puglisi, 2015)⁴⁴.

9.2.1.1.4 Definitions of Political Bias in News

The review found three examples of specific lexical definitions, descriptions or conceptualisations of political bias in news in the literature. The earliest is that of Stevenson et al. who reran Efron’s earlier study. They noted several basic shortcomings in Efron’s work, including methodological issues. As they state in their introduction: *“...the examples of ‘bias’ which she noted were at best questionable.”* (Efron, 1971; Stevenson et al., 1973). Their study, titled “Untwisting the News Twisters: A Replication of Efron’s Study” was designed to address these issues. Like Efron, they did not put forward a specific definition of bias, but offer their conceptualisation of the term, deciding to stay close to Efron’s conceptualisation, see section 9.2.1.1.3. *“Finally, we decided to conceptualize bias as imbalance or inequality of coverage rather than as a departure from truth. While the latter definition can be used in some circumstances, it did not seem possible or useful to try to develop a ‘true’ picture of the campaign to compare with CBS’ version. And the definition of bias as imbalance had the advantage of being used both by Efron and the Federal Communications Commission in its ‘fairness doctrine.’”* While this is a conceptualisation rather than a specific lexical definition of the construct, it is clearer than Efron’s definition by understanding of the FCC Fairness Doctrine. It should be noted that despite its shortcomings and detractors, Efron’s work is a noteworthy study in the domain, being not only one of the earliest, but also for initially recognising the difficulty of the task and her attempt at a solution, and thus deserves its place in this review.

⁴⁴ These findings are quite pertinent and topical considering the Irish Government’s short lived Strategic Communications Unit (SCU) which was set up with a budget of €5m per year in 2017 (Callan, 2017). Significant controversy has surrounded this unit since its inception. It was tasked with communicating government success stories through the media (Noel Whelan, 2018; Ó Cionnaith et al., 2018; The Irish Times View, 2018). The unit was found to be purchasing advertising in often financially hard pressed local newspapers to communicate good news about the government’s new ‘Ireland 2040 National Planning Framework’. These advertisements were in many cases deliberately disguised as news articles in the form of ‘advertorials’ which the SCU had final copy approval of (Coyne, 2018a, 2018b, 2018c). Subsequently news reports surfaced that the local news agencies were essentially being pressured to accept the SCU’s demands for oversight of the means, placement, styling, and final content approval with the promise of more such advertisements in future. As a result of the controversy and accusations of propaganda by the opposition in the Dáil (national parliament) the SCU is currently under review and will likely be closed down (C. Finn, 2018; Lehane, 2018; McEnroe et al., 2018).

An explicit lexical definition of political bias in news was that put forward by Kenney and Simpson with: *“Bias is defined as a pattern of constant favoritism. Bias occurs when one candidate or party receives more news coverage and more favorable coverage over an extended period of time.”* (Kenney & Simpson, 1993). Their study of content bias of the 1998 US election demonstrated that overall, most of the major dailies coverage was unbiased, except for that of the conservative Washington Times which favoured the Republican Party. Their decision to root their definition in ‘favouritism’ contrasted with many earlier studies which focused on balance or omission. This actually has similarities to the conceptualisations of the construct used on studies of racial bias by Rainville and McCormick and later by Rada, which focused on the positive and negative portrayal of African American footballers by sports presenters (Rada, 1996; Rainville & McCormick, 1977).

In their study of political bias in the news coverage of Taiwan’s first presidential election in 1996, Lo et al. adopted Kenny and Simpson’s definition: *“This study follows the simple definition of bias used by Kenney and Simpson (1993, p. 346). Following them, bias is defined as ‘a pattern of constant favouritism’.”* (Lo et al., 1996). Their study demonstrated that state owned television stations gave a greater amount of coverage to the ruling party and that the coverage was more favourable. Lo also adopted this definition in a later study into journalistic bias in news source selection with: *“It [bias] may also refer to ‘systematically favouring one side or a certain stance’”* (Lo, 1995).

9.2.1.1.5 Descriptions of Content or Photograph Bias

There are also several descriptions specific to content bias, be it textual content or photographic content, in the literature. These are often contained in studies with very specific areas of focus and/or which use some form of a scale or model, to measure whether specific items of information had a positive or negative impact. A full description of such empirical methodologies can be seen in D’Angelo and Kuypers book (D’Angelo & Kuypers, 2010). It should be noted that the following studies do not include lexical definitions but rather have descriptions of content bias used in these and similar experiments. They are presented here as an alternative approach and to demonstrate that not all studies use specific definitions. None of these studies put forward a specific definition of bias, but rather measured the amount of positive and negative content in whatever form including non-verbal communication. One possible reason for the lack of a specific definitions for the material being measured, which is rare in peer reviewed academic research, are the aforementioned issues in defining bias.

In a Friedman et al. study of the facial expressions of newscasters in the 1976 presidential campaign, they used a twenty-one-point scale to measure non-verbal communication, specifically the positivity and negativity of newscasters facial expressions after they said a candidate’s name (H. S. Friedman et al., 1980). In Robinson and Sheehan’s book on the 1980 campaign, they measured the level of objectivity, fairness, seriousness, and criticism, displayed by CBS News and United Press International (UPI) between January

and December of 1980 by analysing among other things, the relative amount of positive and negative quotes and comments (M. J. Robinson & Sheehan, 1983). In their study of newsmagazine visuals during the 1998 presidential campaign, Moriarty and Popovich measured the positive or negative statements relating to character, job performance, campaign performance, and issue stands (Moriarty & Popovich, 1991). In their study of bias in soundbites from major network newscasts over two presidential campaigns, Lowry and Shidler coded soundbites from 199 newscasts to measure the amount of positive and negative quotes and comments (D. T. Lowry & Shidler, 1998). All three of Lowry's studies into network television news, including this one, relied on the Hayakawa-Lowry system of categorisation (D. T. Lowry, 1986), which was first established by the linguist Hayakawa (Hayakawa, 1949).

9.2.1.1.6 *Definitions by Differentials in Coverage*

In a departure from previous approaches in the domain, Watanabe highlights issues with existing types of definitions. These include those based on objectivity, which is often measured in terms of positive or negative tones, or framing: *“which become either positive or negative when they concern events, favourable or unfavourable when they concern opinions, or supportive or critical when they concern policy options”* (Watanabe, 2017b). He maintains that there are issues with constructs such as fairness and balance as proxies for objectivity. Consequently, he presents an innovative approach, whereby he: *“conceptualize objectivity in news reporting as coverage of all possible newsworthy stories”*. This was undertaken by comparing the coverage of the 2014 Ukraine crisis and its lead up between Russia's state news agency, The Information Telegraph Agency of Russia, Telegraph Agency of Sovereign States (ITAR-TASS) with Interfax, an independent Russian news agency. Arguing that Interfax is the benchmark for coverage of events, he tracked the coverage bias in ITAR-TASS over a 16-months, from January 2013. The results demonstrate the Russian governments influence on ITAR-TASS, which: *“clearly show that ITAR-TASS' framing of democracy and sovereignty in Ukraine is systematically biased during the crisis corresponding to the desirability of the situation in Ukraine for the Russian regime.”* To uncover this, he focused on both agencies coverage of pivotal events including the positive or negative framing of democracy and sovereignty. He goes on to say: *“The main causes of bias were (1) highly critical comments made by Russian officials on Ukraine, which the news agency quotes very frequently, and (2) profoundly negative descriptions of events related to Ukraine by the news agency.”* The newness of this approach means there is no way to tell whether it will be widely adopted.

9.2.1.1.7 *Categorisation of Definitions of Bias by Related Terms*

Stevenson and Green point out that most definitions of bias fall into two categories, the lesser used is inaccuracy and the more common is selection (Stevenson & Greene, 1980). However, this review has demonstrated that this categorisation is no longer relevant. For this review, definitions of bias have been categorised by the domain they specifically pertain to in their text or in which they were applied. This methodology, while not perfect, makes the review and comparison of definitions easier. It will also make

it a more useful resource for consultation. As bias is a subjective construct, one of the most common methods of defining it is by means of another construct such as truth, or the absence of a specific element such as objectivity. Often these take the form of specific definitions of what it is, and what it is not. Table 9-1 depicts all the other constructs or elements used in the previous definitions to define the construct of bias. While there is variation in the related terms, their relationship is evident when they are considered in relation to bias. Many are dimensions of credibility in their own right. Of course, not all the definitions presented thus far used related terms to describe bias, such as those by (Di Tella & Franceschelli, 2011; Kline, 1981; Philo et al., 1977; Russo, 1971). Many of the definitions highlighted here also make some reference to time which is shown in the centre column.

Table 9-1 Use of relied upon terms to define bias. The centre column denotes whether indications of time, and the term, that have been used in the authors definition.

Related Term	Continuum	Author
Positive / Negative		(Miller et al., 2007; Rada, 1996; Rainville & McCormick, 1977)
Positive / Negative / Neutral * Descriptions /conceptualisations		(H. S. Friedman et al., 1980; D. T. Lowry & Shidler, 1998; Moriarty & Popovich, 1991; M. J. Robinson & Sheehan, 1983)
Truth	Consistent	(McQuail, 1992)
Slant†		(Groseclose & Milyo, 2005a)
Slant	Systematic	(Waldman & Devitt, 1998)
Patterns of slant		(Entman, 2007, p. 166)
Fairness		(Efron, 1971)
Constant favouritism	Constant	(Kenney & Simpson, 1993; Lo et al., 1996)
Systematically favouring		(Lo, 1995)
Selective omission, Choice of words Varying credibility		(Gentzkow & Shapiro, 2006).
Portrayal of reality	Systematically	(Groeling, 2013a)
Opposite of objectivity		(Hoffman, 2013; T.-T. Lee, 2008)
Absence of objectivity		(Doll & Bradley, 1974)
Partiality		(Hofstetter, 1976)
Differential treatment‡	Systematic	(Stevenson & Greene, 1980)
Selective Inaccuracy		(Gaddy & Tanjong, 1986)
Imbalance or inequality		(Stevenson et al., 1973)
Differential		(Klein & Maccoby, 1954; D. T. Lowry, 1973)

* *Descriptions or Conceptualisations, not strictly definitions*

† *Encouraged people to substitute whatever term they want (taste or preference)*

‡ *Argued this was the standard definition and put forward an alternative*

9.2.1.1.8 Algorithmic Definitions of Bias

Many approaches have been used to measure or quantify bias. These range from the use of scales (H. S. Friedman et al., 1980), quantifying statements as for or against a candidate, position or ideology (Moriarty & Popovich, 1991), size or favourability of photos (Barrett & Barrington, 2005), the measuring of column inches in newspapers (Di Tella & Franceschelli, 2011), questions about favourability (Hoffman, 2013), or questions about the absence of objectivity (Doll & Bradley, 1974). There have also been several attempts at developing formulaic methods to measure the construct, often by those with an economics background, who attempt to account supply side or market side forces (Gentzkow et al., 2015). The following section highlights three such attempts to demonstrate that lexical definitions of the term are not always forthcoming and that the construct has attracted serious scholarly endeavour from diverse backgrounds.

Mullainathan & Shleifer used their training as economists to develop a model of how competition in the market affects media bias (Mullainathan & Shleifer, 2005). Their work was undertaken under two suppositions; that readers hold beliefs which they like to see confirmed; and that newspapers can slant stories towards these beliefs. Their work found that competition in the market is not a force for accuracy and in fact, competition forces newspapers to cater to the beliefs of their readers, which may not be an accurate portrayal of events. They also found that on topics where readers' positions diverge, newspapers segment the market and tailor their news towards the beliefs of their desired audience. This is based on the knowledge that consumers find news which agrees with their positions to be of higher quality (Gentzkow & Shapiro, 2006). Mullainathan & Shleifer do note however, that it is possible to gain an unbiased view of the news by the consumption of news from a variety of news outlets. Their algorithm defines bias: "*as the average amount by which the news read deviates from the data for the average reader*".

Entman proposes definitions and improved measures for the linked but separate terms of slant and bias in his attempt to provide an insight into how the media influence the distribution of power (Entman, 2007). He differentiates slant from bias, in that slant is characterised by favouring one side over the other in individual news reports: "*framing favors one side over the other in a current or potential dispute.*" While bias is a consistent pattern of slant or framing over a prolonged period: "*consistent patterns in the framing of mediated communication that promote the influence of one side in conflicts over the use of government power.*" Based on these definitions, he puts forward 'The Aggregate Slant News Index' (ANSI) formula and 'The Bias Index' (BI) formula which uses the ANSI to measure the phenomenon.

Gentzkow, Shapiro and Stone also use their background in economics to develop a model of media bias in the marketplace (Gentzkow et al., 2015). In their contribution, they: "*define bias as a partial order in reporting strategies*". This adopts the definition put forward in their earlier work detailed above (Gentzkow & Shapiro, 2006). They go on to say: "*Under our definition, bias is only defined as a relative concept: one outlet can be to the right of another, but only in special cases is it meaningful to talk about 'unbiased' or 'objective' reporting, about 'more' or 'less' bias, or about whether a given outlet is left or right of 'center.'*" The subsequent model they propose allows them to account for the determinants of bias, highlighting that these can be both supply side, such as the journalist, editor or news agency, or demand side, the consumer.

9.2.1.1.9 Definitions of Bias in Social Science

One definition that attracted our attention which is not from a journalistic, information science or even an economics background, is that of Hammersley and Gomm who proposed a conceptual framework for recognising bias in Social Science research (Hammersley & Gomm, 1997). They define "*'bias' as one of*

several potential forms of error” and most importantly differentiate between motivated and unmotivated and conscious and unconscious, see Figure 9-1.

They go on to say “*Within this framework we could define ‘bias’ in several different ways. We might, for example, restrict its meaning to systematic, culpable, motivated, error. Alternatively, we could treat this as one form of bias, using the term to refer to all kinds of culpable, systematic error, or even to all kinds of systematic error. There seems little advantage to defining it as all systematic error, since this involves a duplication of terms, and there are other important distinctions to be made. Our own preference is to define ‘bias’ as systematic and culpable error; systematic error that the researcher should have been able to recognize and minimize, as judged either by the researcher him or herself (in retrospect) or by others. This then allows us to distinguish between motivated and unmotivated bias, according to whether it stems from other goals than the pursuit of knowledge*” (emphasis added after) (Hammersley & Gomm, 1997). This definition is interesting as a comparison as it is the only attempt to define bias as fundamentally, erroneous reporting of the facts. By replacing the word ‘researcher’ with ‘journalist’ in the above paragraph we can see the difference between the domains. In this definition, the culpability is put directly on the originator who should be capable of owning up to their mistake. In comparison, none of the aforementioned definitions took such an approach.

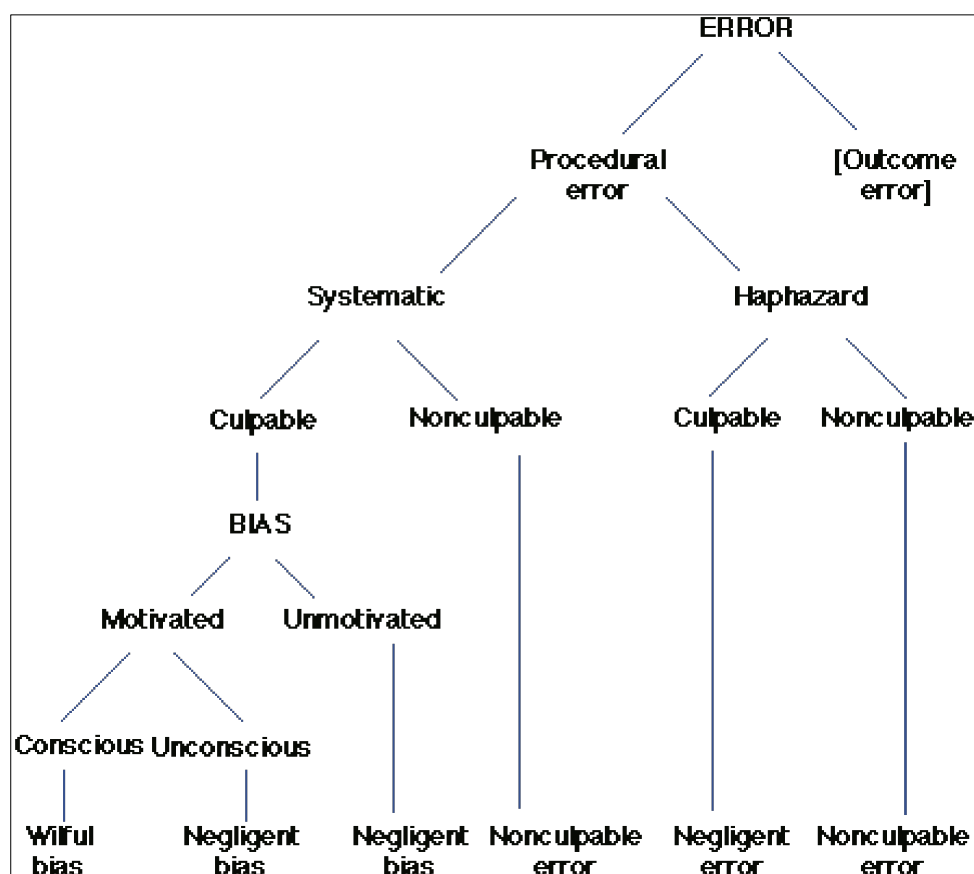


Figure 9-1 A conceptual framework by Hammersley and Gomm which categorises different types of bias in social science research as different types of error (Hammersley and Gomm, 1997).

9.3 Appendix C – Chapter Two: Biases in the Production of News

The following continues the section 2.3.4.1 “Bias in the Production of News”. While it was out of scope for the main thesis document it is still relevant to the research as a whole, and to demonstrate that bias in the production of news had been fully researched.

A paper by Park et al. entitled, “NewsCube: Delivering Multiple Aspects of News to Mitigate Media Bias” (Park et al., 2009) designed a service to present multiple articles surrounding each event from different news outlets, thus providing the consumer with a plural of viewpoints to help mitigate bias. It should be noted that one of the findings from the Mullainathan and Shleifer contribution was that it was possible for consumers to gain unbiased news from an aggregate of news sources (Mullainathan & Shleifer, 2005). A noteworthy and oft unnoticed contribution of the paper is the graphical explanation, presented in Figure 9-2, of the means by which bias can enter the news production process, the first of the three stages of the news pipeline, production, dissemination, and consumption.

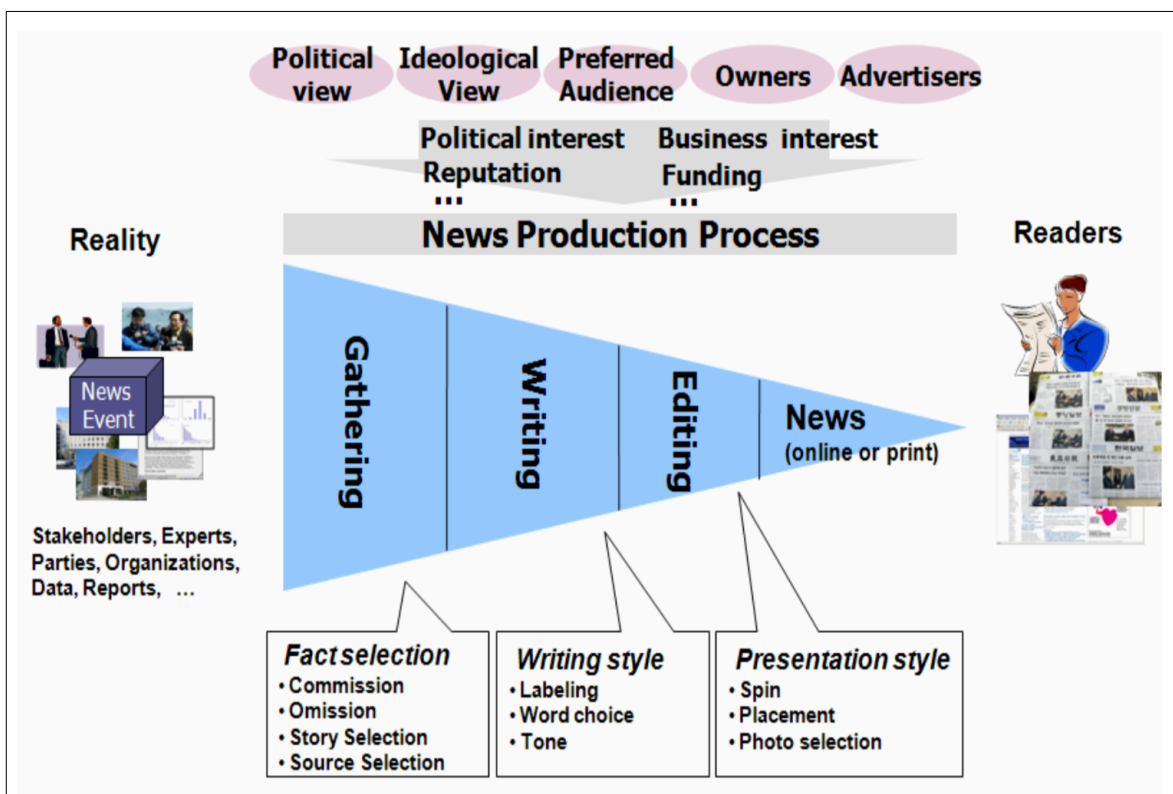


Figure 9-2 Causes and Forms of Media Bias, from (Park et al. 2009).

It breaks down the production stage into three sub-stages; gathering, writing, and editing. At each stage, they: “map the forms of bias embodiment”. At the gathering stage, it includes commission, omission, story selection, and source selection biases. During the writing stage, it includes labelling, word choice, and tone. While during the editing stage it includes spin, placement, and photo selection. At the top of the graphic

the outside influences on the news production process are shown. They include political views, ideological views, preferred audience, owners and advertisers. These form the political interest and the business interest. Each of these may influence or bias the production of news. While this is an important contribution, it is too abstract and does not detail the individual biases that are studied in the domain such as negativity bias and/or the liberal media bias. It also has no consideration for journalistic biases, all of which have been studied extensively. In fact, no complete classification could be found in the literature to explain the depth and complexity of the impact of various biases on any stage of the news pipeline, production, dissemination or consumption.

The following sections represent the first attempt at a categorisation of biases that impact the production stages of the news pipeline. It is heavily influenced by the graphic of Park et al. in Figure 9-2, but it is designed to significantly expand upon it. The fact that this table is currently the most complete categorisation demonstrates the piecemeal state of affairs of the domain and the massive taxonomical task to provide clarity. These tables are not complete. They were put together as the best means of representing the domain as it currently stands. The aim is to extend and publicise them in future. The types of production biases were derived from the most common forms of bias studied and reported on in the domain. It also highlights relevant papers which specifically deal with each form of bias. It should also be noted that there is huge area of overlap between many of the forms of bias and the papers that report on them. One reason for this is the integrated and overlapping nature of many of the forms of bias, e.g., selection bias and partisan bias, another is the fact that some authors use terms such as framing and agenda setting interchangeably.

The following sections detail some of the Global, Domain, Organisational, Editorial and Journalistic biases that influence the news production stages of; commissioning, fact and information gathering, writing, and editing. These are presented as a brief overview of a complex issue to provide context. However, they are not the focus of this research, which is focused on bias in the dissemination of news. The format of the following tables is influenced by the format of the table 'Website quality dimensions across studies' by Kim and Stoel, and by the table 'Online attributes investigated by various scholars', which describes features and characteristics of ecommerce websites which have been investigated for their effect on website quality, by Wolfinbarger and Gily (S. Kim & Stoel, 2004b; Wolfinbarger & Gilly, 2003).

9.3.1 Global Influences that Bias the Production of News

The most prevalent global bias issue is that of government influence on the press. This is most evident in the divide between the relatively free western press and the press in predominantly developing countries which in many cases has some element of government control. Instances of concern include China, Russia,

Saudi Arabia, and Iran and many other countries in the Middle East⁴⁵ where journalists often risk considerable pressure or worse from the government⁴⁶. In many instances the press is pleased or influenced directly by government to produce news with a certain slant. In Egypt alone there have been several recent incidents of journalists from international organisations being detained⁴⁷⁻⁴⁸, A free press has also been shown to be a deterrent to corruption (Brunetti & Weder, 2003). Another domain concern is the liberal conservative divide in the media, with many especially in the conservative echelons of society, harbouring or even stoking a fear that too much of the press is liberal. Several contributions support this claim including that of Groseclose who found that most mainstream media outlets have a liberal bias, and even so called conservative media outlets, lean less right than their left leaning counterparts (Groseclose, 2011). Some of the global influences which may bias the production of news are presented in Table 9-2.

Table 9-2 Global influences which may bias the production of news.

Global Influences on the Production of News	Manifestation / Effect	Publication
Government	Reduced coverage of political scandal, positive or negative reporting of government policy, law, event, or figure	(Di Tella & Franceschelli, 2011; Watanabe, 2017b)
Ideological	Partisan reporting to promote one candidate, party or ideology	(Waldman & Devitt, 1998) (Groseclose, 2011)
Nationalistic	Nationalistic decisions by referees in sporting events. Partisan reporting in the news to favour a country or population often during times of conflict or heightened tension	(Billings & Tambosi, 2004; Emerson et al., 2009; Horvit, 2006; Montiel et al., 2014)
Religious	Religious influence on the production of news	(Baker, 2010; Pintak, 2014; Ranly, 1979; Scott & McDonald, 2005)

9.3.2 Domain Influences that Bias the Production of News

There are several known domain influences that commonly bias news. The most obvious is the difference between tabloid and quality press. Serious newspapers tend to focus on national, international or government issues in the preceding weeks and weeks ahead. Due to the nature of their format they tend to offer more analysis of events. In comparison, daily tabloids are more reactionary and therefore tend to focus on daily events and human-interest stories. The medium may also bias the news produced and disseminated by a news organisation. An example is the dichotomy between talk radio and late-night news shows in the US. Conservatives embraced talk radio due to the lower socioeconomic status of much of its audience, which primarily consisted of farmers, factory workers, etc. While liberals embraced late night

⁴⁵ Reporters Without Borders World Press Freedom Index 2017 https://rsf.org/en/ranking_table and a multi-year table showing data from the same organisation https://en.wikipedia.org/wiki/Press_Freedom_Index

⁴⁶ International Federation of Journalists <http://ifj.org/> tracks journalist and media staff deaths around the world. In 2017, 82 lost their lives <http://ifj-safety.org/en/2017/killings>

⁴⁷ At the time of writing, (31/03/ 2019) Al Jazeera Journalist Mahmoud Hussein has been detained for 828 days without charge. He was arrested while on holiday in Egypt and accused of “incitement against state institutions and broadcasting false news with the aim of spreading chaos” <https://www.aljazeera.com/programmes/aljazeera-world/2018/02/journalism-crime-story-mahmoud-hussein-180212140543338.html>

⁴⁸ Three Al Jazeera journalists given seven-year jail sentences in 2014 for spreading false news <http://www.bbc.com/news/world-middle-east-27943387>

news talk shows as their preferred format. Some of the domain influences which may bias the production of news are presented in Table 9-3.

Table 9-3 Domain influences which may bias the production of news.

Domain Influences on the Production of News	Manifestation / Effect	Publication
Transnational v National v Local	Different treatments of the same topic, different focus based on audience	(E. Kim & Lowrey, 2018; H. Li & Tang, 2009)
Medium / Format	Conservative talk radio v Liberal late-night talk shows	(Hoffman, 2013)
Tabloid / Quality	Different representations of race, issues, religion or political parties	(Van Dijk, 2000; Vergeer et al., 2000)

9.3.3 Organisational Influences that Bias the Production of News

As Park et al, so astutely point out, organisational bias can be influenced by factors including political, ideological, preferred audience, owner, and/or advertisers (Park et al., 2009). Since the 1970s successive studies have demonstrated that news organisations, whether newspapers agencies, television networks, and to a lesser extent radio stations, are prone to bias in one form or another (Stefano DellaVigna & Kaplan, 2007; Efron, 1971; Groeling, 2008, 2013a; Groseclose & Milyo, 2005a; Hofstetter, 1976; Kenney & Simpson, 1993; Lo et al., 1996). Examples of political bias includes the litany of studies showing that different networks and newspapers favour one candidate or party in US presidential elections (Batlin, 1954; Diddi et al., 2014; Groeling & Kernell, 1998; Klein & Maccoby, 1954; Mantler & Whiteman, 1995). Ideological bias in news organisation was expertly dealt with by Bovits et al. who asked the question: “*Do news organisations purposefully lead the public to support a particular ideological agenda?*” (Bovitz et al., 2002). Their model demonstrates when politically driven media magnates can achieve or support political goals and ideological causes in combination with market forces. This issue has been flagged for at least a century. A royal commission into the press from 1947-1949 demanded by journalist and some Labour politicians based upon the view that: “*powerful chains formed a monopoly tending to stifle independent and local publications and that financial and advertising interests, combined with a concentration of press control in the hands of a few publishers, was resulting in the distortion and suppression of information.*” (R. W. Desmond, 1949; *Royal Commission on the Press, 1947-1949*, 1949).

As previously elucidated, significant work has been done to show that many media organisations deliberately slant their news to suit their desired audiences (Gentzkow & Shapiro, 2010; Mullainathan & Shleifer, 2005). While this had been suspected and even reported for some time, its demonstration was an important milestone in the domain. It is also worth noting that the graphic from Park et al. also depicts ‘Preferred Audience’ as a specific influence on bias, that Figure 9-2, (Park et al., 2009). More recently Gentzkow, Shapiro and Stone theorise on the market influences on bias in news organisations and differentiate between supply side bias and demand side bias (Gentzkow et al., 2015). The work of Iyengar

and Hahn has since demonstrated that consumers exercise selective exposure to news, based on anticipated agreement with its content (Iyengar & Hahn, 2009).

Even a surface level review of the literature shows that the amount of influence advertisers can bring to bear on journalists seems heavily dependent on how reliant they are on them as source of income. Thus, local and niche publications suffer worse than national publications, though they are not immune either. In a survey of 190 agricultural journalists in local or niche publications, Hays and Reisner show that advertisers try to make sure that copy is favourable to their product. 62% of journalists reported receiving some pressure while approximately half maintain that advertising had been withdrawn due to unsupportive practices (Hays & Reisner, 1991). In a survey of network news correspondents in the US to determine if owners or advertisers influence news, 20% of participants felt some ownership pressure to report or censor news, while 7% felt advertiser pressure (Price, 2003). While this is relatively low, when 219 advertising sales executives in daily newspapers were directly surveyed, they reported frequent conflicts between the business operations and journalistic practices (An & Bergen, 2007). The much higher incidents in An and Bergen's study is likely due to the fact that journalists do not often interact with advertisers directly whereas the sales team do. As demonstrated earlier, the impact of purchasing advertising, both by government and commercial organisations on a free press has been shown in markets such as Argentina (Di Tella & Franceschelli, 2011) and Italy (Gambaro & Puglisi, 2015). In a large study with 291 fashion companies and 123 publishers, Rinallo and Basuroy found: "*evidence of a strong positive influence of advertising on coverage*" (Rinallo & Basuroy, 2009). Further demonstrating that local, niche or special interest magazines are more susceptible to the influence of advertisers than national press, Reuter and Zitzewitz found that mutual fund recommendations correlate with past advertising in three personal finance publications but not in national newspapers (Reuter & Zitzewitz, 2006).

Multiple studies have been conducted comparing and contracting news organisations in the US, with most highlighting the conservative nature of Fox news, The Drudge Report and/or the Washington Times and the predominately liberal nature of the NBC, ABC, CBS, CNN, NYT and the Washington Post (Gentzkow & Shapiro, 2010; Groeling, 2008; Groseclose & Milyo, 2005a). The conservative Fox news in particular has been studied extensively. Morris focuses on the increasingly polarisation of Fox news and CNN audiences, showing that Fox news viewers were less likely to follow news critical of the Bush administration and that they were more likely to follow news that shares their personal view (Morris, 2005). The highly cited and influential study by DellaVigna and Kaplan demonstrated the impact of the introduction of the conservative Fox news to an area and increased conservative voting patterns. Using a dataset of 9,256 towns, they showed that Republicans gained 0.4 – 0.7 in vote share when Fox news was introduced. The analysis also suggests that Fox news convinced 3% – 28% of its audience to vote Republican. Conversely, Gerber et al. demonstrated that introducing newspapers subscriptions before an election to either the liberal Washing Post or the conservative Washington Times led to an increase in

support for the Democratic candidate. Their study suggests that media slant matters less than media exposure (Gerber et al., 2009). This should be understood in the context of recent work showing that news agencies bias their news in favour of the audiences positions, especially on polarising topics (Gentzkow & Shapiro, 2006; Mullainathan & Shleifer, 2005). It must be noted that anecdotal evidence and even a surface level review of the literature suggest that the so-called liberal media in the US, especially the impact they may have on consumers' viewpoints and voting habits does not face the same level of scrutiny from academia.

In comparison, the BBC as a public service broadcaster, it is mandated to provide fair and impartial coverage. It is perhaps this fact that has resulted in both liberals and conservatives accusing it of bias in favour of the other side. Though it has historically been seen as slightly more conservative in nature (Ayton & Tumber, 2001). Perhaps in a sign of understanding which way the wind was blowing, the perception of conservative bias reduced dramatically with the election of the Blair Labour government in 1997. It must be noted however that despite these claims and counter claims and the timely switch in coverage favouring the government, the BBC is regularly considered one of the least biased news agencies. This was confirmed in a report on their coverage of the second Intifada instigated by pro-Israeli lobbyists who warned of an anti-Israel bias. The study found that if anything, the BBC was more pro-Israel in its coverage (Gaber et al., 2009). Several studies have found that overall the BBC is mostly unbiased in its liberal / conservative coverage, though recently there is a left of centre tendency. This was most recently demonstrated by Latham who used a content analysis to identify citations to think tanks and warnings in the news article text about their affiliations and/or independence (Latham, 2013). Groseclose and Milyo pioneered this approach in the US (Groseclose & Milyo, 2005a). Krestel et al. also used a similar approach (Krestel et al., 2012).

Organisational news bias is a serious concern in the domain and should be considered in any study of bias in the news. As Park et al. demonstrates in Figure 9-2, organisational bias can influence the gathering, writing and editing stages of news production, the first stage in the news pipeline (Park et al., 2009). The above paragraphs highlight some of the most pertinent studies on organisational bias in the literature. However, news organisation bias is not the focus of this thesis and they are only provided for context. Some of the organisational influences which may bias the production of news are categorised in Table 9-4.

Table 9-4 Organisational influences which may bias the production of news.

Organisational Influences on the Production of News	Manifestation / Effect	Publication
Owners	News in favour of the owners' position on a topic or ideology. Examples include corporate regulation, lobbying, appointment of judges, and environmental concerns	(D. W. D'Alessio & Allen, 2007; Gilens & Hertzman, 2000; S. Hughes & Lawson, 2004; Price, 2003)
Political	Reduced or increased coverage of a politician, party, or scandal based on pressure from the government, a political party, or political movement,	(Di Tella & Franceschelli, 2011)
Ideological	Examples include: Left: Socialism, Communism, Marxism, Leninism Right; Fascism, Nazism, anti-immigrant Conservative v Liberal Business / Industry v Environmental / Green Pro-Government v Anti-Government Pro-business / Free Market v Pro Worker / Planned Market Religious Conservative v Social Progressive	(Bakshy et al., 2015; Bovitz et al., 2002; Eisinger et al., 2007; Iyengar & Hahn, 2009; Turner, 2007)
Preferred Audience	Slanting news in a manner to appeal with their preferred audiences' positions on a range of issues. This is based on the knowledge that audiences find news that agrees with their position to be more credible and are therefore more likely to consume it	(Gentzkow et al., 2015; Gentzkow & Shapiro, 2010; Mullainathan & Shleifer, 2005)
Commercial / Advertisers	Reduced or increased coverage of a product, service or scandal involving a commercial partner or advertiser	(An & Bergen, 2007; Di Tella & Franceschelli, 2011; Gambaro & Puglisi, 2015; Hays & Reisner, 1991; Price, 2003; Reuter & Zitzewitz, 2006; Rinaldo & Basuroy, 2009; Shoemaker & Reese, 1996)

9.3.4 Journalistic or Editorial Biases that Influence the Production of News

Journalists have been studied as a source of bias since at least the 1960s. Kerrick et al. demonstrated that journalistic students bias articles in favour of a given editorial policy and most interestingly, the student who disagrees most with the policy writes the most partisan content in favour of the policy (Kerrick et al., 1964). As part of a series into reporting of elections in Taiwan, Lo demonstrated that journalists are biased in their selection of sources (Lo, 1995). In a survey of journalists in five countries, Patterson and Donsbagh revealed a significant correlation between journalists personal beliefs and their news decisions and that this was more pronounced among newspaper journalists than broadcast journalists (Patterson & Donsbagh, 1996). A study of source bias in Canadian television news showed that journalists selected politicians or their representatives from government as sources rather than the opposition, which allowed them to shape the debate. This was also true when Canadian journalists selected foreign politicians as sources (Hackett, 1985). In their proposal for the study of cognitive biases in journalists, Stocking and Gross list out a number of factors which bias journalists including; deadlines, organisational structure, and a: *“host of other environmental factors”*. They cite Nord in their claim that: *“If news is biased, it is argued, it has less to do with the individuals who process the news and more to do with the organizations, communities, and cultures within which they work (Nord, 1985)”* However, it must be noted that this claim has received little

support in the domain and while Stocking and Gross do not dismiss environmental factors, they maintain that cognitive forces have more influence (Nord, 1985; Stocking & Gross, 1989). Rucinski adeptly highlighted the practices of some journalists who bias their reporting with their tendency to focus on, or create a personal angle to stories (Rucinski, 1992). More recently Ekström et al. demonstrated partisan bias in journalists during a review of politician interviews over three elections in Sweden (Ekström et al., 2013). The latest contribution in the domain strikes a note of caution, and has a somewhat ironic note to it. In a survey conducted on journalists during the Iraq war, 17.5% of journalists embedded with troops said that the stories from embedded journalists were biased while 54.6% said they were not biased. Presumably 27.9% were neutral on the question (Fahmy & Johnson, 2005). It should be noted that in an earlier study by Pfau et al. on the impact on tone and framing of embedding journalists in military combat units during the Iraq War, that reporting by embedded print journalists was more favourable in overall tone towards the military and towards the troops (Pfau et al., 2004).

A range of studies have been undertaken on editorial policy and influence on bias. This is based on the idea that editors in the newsroom, and in their editorials, position a paper for or against a politician, policy, party or ideology etc., and thus employ or encourage journalists to take a similar stand. Or as some have suggested, editorial policy is the inkwell from which bias is spilled over the paper (Forward, 1977; C. A. Hughes & Western, 1973). Editors can influence news in a number of ways, including by explicit means through their editorials and hidden means such as gatekeeping, journalist assignment, post editing, and photo selection among others. Hackett provides an exemplary discussion of editorial news bias (Hackett, 1984). One of the most powerful forms of editorial bias is that of gatekeeping. In one of the formative studies in the domain titled “The ‘Gate Keeper’: A Case Study in the Selection of News” and the first study of its type, White recruited a wire editor of a morning newspaper with a circulation of approximately 30,000 (White, 1950). The editor was in his mid 40s with 25 years of experience in the domain. His name was withheld but he was given the moniker “*Mr Gates*” for the study. Mr Gates was asked to keep a log of all the wire copy he received including that which he rejected. At the end of the evening he had to note the reason why each piece of wire copy was rejected if he could remember. The study found that 90% of copy was rejected and that the reasons were in many cases highly subjective. There were eight categories of reasons for rejecting the copy as worthy of reporting. Three of them highlighting the subjective nature of the rejection are highlighted below.

- *“Trivial (29); would ignore (21); no need for this; wasted space; not too important; not too hot; not too worthy”*
- *“Propaganda (16); he’s too red; sour grapes”*
- *“Wouldn’t use (11); don’t care for suicide stories; too suggestive; out of good taste”*

The 'Gate keeping' study of its kind and clearly demonstrate that the news we experience is selected for us and represents only a small proportion of the news. Gatekeeping is now one of the core theories of communications and journalism studies (Shoemaker & Vos, 2009).

Editorial policy is another form by which editors may influence or bias news. As previously stated, Kerrick found that journalistic students would adhere to editorial policy by deliberately biasing their articles to abide by it, and most interestingly this effect was strongest in those who disagreed with the policy most (Kerrick et al., 1964). Forward also found some evidence which suggests editorial opinion influences journalistic writing. In a survey of 9 Australian papers over 10 months, he found that only 21% out of 2,006 editorial items were positive while 42% were negative. In his analysis of the content of the same papers he found that they were hostile to the government 76.7% of the time. This suggests that editorial policy has some influence on journalist's contributions. The study also found that editorial policy closely reflected public opinion and possibly even pre-empted or shaped it (Forward, 1977).

The influence of editors on journalists is by no means clear cut. In a study of editorial endorsements in 26 newspapers in Wisconsin on coverage of a financial controversy surrounding a vice presidential candidate, Merron and Gaddy maintain that: "*If editorial perspective is an influence on news play, it doesn't appear to be a strong one.*" (Merron & Gaddy, 1986). One of the main accusations by conservatives against the editorial policy of 'liberal media' is that they routinely label conservative politicians as conservative but do not label liberal politicians as liberal. Eisinger et al. found that disproportionate labelling of conservatives exists, however they state that this is not done: "*in a way that constitutes 'bias'*" as newspapers often label liberals too (Eisinger et al., 2007).

In a possible sign of the changing news landscape, two studies demonstrate the impact of editorials as harbingers of bias in the content of news articles. Kahn and Kenney investigated the impact of editorial endorsements on campaign coverage and citizens views of candidates (Kahn & Kenney, 2002). Their analysis raises concerns over what they label: "*the wall of separation*", the divide between the political beliefs of owners and editors as expressed in editorials and the news pages of a paper. In a study of 60 US senatorial campaigns in three election years, they found that information on news pages is slanted in favour of candidates supported or endorsed in the editorial pages. The impact of editorial slant on voters was subsequently investigated by Druckman and Parkin (Druckman & Parkin, 2005). They found compelling evidence that editorial slant: "*defined as the quantity and tone of a newspaper's candidate coverage as influenced by its editorial position*", influenced voters' decisions.

Some of the journalistic and editorial biases that may influence the production of news are provided in Table 9-5.

Table 9-5 Journalistic or editorial influences which may bias the production of news.

Journalistic or Editorial Influences on the Production of News	Manifestation / Effect	Publication
Editorial Gatekeeping	Selection of news for consumption by the public, typically by the editor. This has a large overlap with Story and Photo Selection below	(Cassidy, 2006; D. D'Alessio & Allen, 2000; Groeling, 2013a; Shoemaker & Vos, 2009; White, 1950)
Coverage	The amount of coverage of a topic, issue, political party or individual	(Brandenburg, 2005, 2006; D. D'Alessio & Allen, 2000; S. Kingsbury & Hart, 1933; Klein & Maccoby, 1954; D. T. Lowry, 1973; Nokelainen & Kannainen, 2018; Puglisi & Snyder, 2011)decisions
Statement	The use of more favourable statements about or from one candidate, individual, party or ideology	(Brandenburg, 2005, 2006; D. D'Alessio & Allen, 2000; Krestel et al., 2012)
Story and Photo Selection	The rejection or selection of stories to cover, or the decision to dedicate less or more time, space or coverage to a particular story or topic, by a journalist. This has a large overlap with Editorial Gatekeeping above	(Barrett & Barrington, 2005; Earl et al., 2004; Groeling, 2008; Martin, 1988; McCarthy et al., 1996; Oliver & Maney, 2000; Stocking & Gross, 1989)
Source Selection	The selection of sources by journalist for a story which may provide them with the opportunity to partisan a debate by allowing them to set the agenda or to frame the debate	(Y. Kim & Bae, 2006; W.-Y. Lin et al., 2011; Lo, 1995; Tanner, 2004)
Ideological	Often involves partisan, framing and/or agenda setting bias in favour of one ideology. This may be political, religious, economic, and/or national etc.	(Bovitz et al., 2002; Eisinger et al., 2007; Groseclose & Milyo, 2005b)
Framing	The selection or omission of salient points to assemble a narrative, or the provision of a favourable or unfavourable analysis of a situation, to promote a particular interpretation of an individual, event or viewpoint	(Abraham & Appiah, 2006; Bernstein, 2004; Crandall et al., 2009; D'Angelo & Kuypers, 2010; Entman, 1993, 2007; Grabe & Bucy, 2009; Guzman, 2016; Kuypers, 2002; H. Li & Tang, 2009; D. T. Lowry, 2008; Watanabe, 2017b)
Agenda Setting	Highlighting or providing selective coverage of events to influence their importance in order to set the agenda for discussion in the public domain	(Brandenburg, 2005, 2006; Entman, 2007; Larcinese et al., 2011; McCombs & Shaw, 1972; McMillan, 2005; Mrogers & Wdearing, 1988; Sørensen, 2013)
Partisan	Supporting, either overtly or covertly, one side in a debate by choosing topics, facts or sources	(Aitkin, 1972; Bartels, 2002; Brandenburg, 2006; Dalton et al., 1998; Gentzkow & Shapiro, 2010; Groeling, 2008, 2013a; Larcinese et al., 2011; T.-T. Lee, 2008, 2010; Niven, 1999, 2002; Patterson & Donsbagh, 1996; Schiffer, 2006)
Pro Capitalist, Pro Free Market	A claimed intrinsic pro-business or pro market bias that exists in the news	(Brohman, 1995; Page & Shapiro, 2010; J. Smith et al., 2001)
Anti-Socialism, Anti-Communist	A claimed anti-communism and/or anti-socialism bias in the mainstream media due to their being part of the establishment or being influenced by the government	(Belin, 2002; Burgess, 1997; Ceplair, 2008; Page & Shapiro, 2010)
Minimal Government, Anti-Government, anti-establishment	The tendency to communicate news or views which is anti-establishment, often by casting them as the enemy or as those seeking greater control	(Esser et al., 2016; Page & Shapiro, 2010; Watanabe, 2017a)
Nationalist, Xenophobic, Anti-Immigrant	Typically, a right-wing bias which promotes pro government messages of national interests and identity to the detriment of other nations or communities. Often present or studied in times of conflict or in sporting events commentary	(Billings & Tambosi, 2004; Brennan, 2016; Emerson et al., 2009; Horvit, 2006; Liu, 2006; Montiel et al., 2014; Page & Shapiro, 2010; Sabo et al., 1996)
Pro / Anti Incumbent	New reporting in favour or against the incumbent in an election	(Ansolabehere et al., 2006; Green-Pedersen et al., 2017; D. T. Lowry & Shidler, 1998; Mughan, 1996; Page & Shapiro, 2010; Schiffer, 2006)

Status Quo	The belief that mainstream news organisations are agents of power or of the establishment and that they have a vested interest in presenting news that maintains the status quo	(Bolsen et al., 2014; Crandall et al., 2009; McLeod & Detenber, 1999; Page & Shapiro, 2010)
Frontrunner	The tendency to use more favourable photographs of candidates who are doing well in polls or who have done well in debates	(Waldman & Devitt, 1998)
Class	The tendency of news organisations to produce news for higher income groups as they are more likely to be news consumers and because advertisers prefer that demographic	(Entman, 1990; S. Friedman, 2011; McChesney, 2003)
Negativity, Bad News	The tendency of news organisation to focus on bad or negative news stories or to report on a story from a negative point of view	(Aday, 2010; Brodie et al., 1998; Lengauer et al., 2012; Soroka & McAdams, 2015)
Race	Biased reporting of certain races or ethnic groups. This is often investigated in terms of crime figures, crime reporting or sport	(Entman, 1990; Hofstetter, 1979; McMillan, 2005; Rada, 1996; Rainville & McCormick, 1977; Romer et al., 1998; Van Dijk, 2000; Vergeer et al., 2000)
Gender	The disproportionate coverage of one gender or another, or by one gender or another, in a favourable or unfavourable manner. It may also include the differential treatment based on gender	(R. Desmond & Danilewicz, 2010; Eastman & Billings, 2000; Grabe et al., 2006; Navia & Osorio, 2015; C. Robinson, 2016; K. B. Smith, 1997; Tyler Eastman, 2001)
Local, Proximity, Geographic	The reporting or highlighting of news based on location and the ignoring of far off more pertinent events	(Dominick, 1977; Gaddy & Tanjong, 1986; Gilliam et al., 2002; C. A. Hughes & Western, 1973; Jones, 2008; W.-Y. Lin et al., 2011; Martin, 1988; Van Belle, 2000; Whitney et al., 1989)
Omission	The act of leaving out, hiding, or reducing the prominence of salient facts, whether knowingly or unknowingly, from a news report. It is highlighted in several studies (Groseclose & Milyo, 2005a). but has received very little empirical study	(Stefano DellaVigna & Kennedy, 2011; Wolton, 2017)
Commission	The opposite of omission, it is the phenomenon of deliberately biasing news or other content in favour of a particular position or ideology. Rarely studied, it is highlighted as a means to bias news by some prominent researchers (Groseclose & Milyo, 2005c; Park, Ko, Liu, et al., 2011; C. Robinson, 2016)	(R. F. Carter & Greenberg, 1965; Stefano DellaVigna & Kennedy, 2011; Groseclose & Milyo, 2005a; Jiwani, 1993).
Labelling	The use of labels, often in a pejorative manner, to identify someone or something as part of a group whose ideology may be at odds with the author and their readership e.g. “The liberal congressman”, “The conservative think-tank” or “The abortion supporting Judge”	(Eisinger et al., 2007; Goldberg, 2003)
Endorsement / Opposition	The belief that journalists, editors, or news organisations would give more favourable coverage to candidates or parties who they have endorsed and less favourable coverage to those they oppose	(Ansolabehere et al., 2006; Chiang & Knight, 2011; Druckman & Parkin, 2005; Kahn & Kenney, 2002; Merron & Gaddy, 1986; Puglisi & Snyder, 2011; Shoemaker & Reese, 1996; Wilhoit & Auh, 1974)
Selective Perception	The phenomenon of not noticing and/or more quickly forgetting stimuli that contradicts prior beliefs or causes discomfort	(Donsbach, 2004; T. P. Meyer, 1972)
Description	The veracity of the coverage of events by the description given to them	(Earl et al., 2004; McPhail & Schweingruber, 1998; Owens & Palmer, 2003; J. Smith et al., 2001)
Intergroup, Stereotype	Linked to stereotype bias, intergroup bias is a systematic bias in language often used to promote stereotypes and separate groups. It usually promotes positive behaviour by in group members and negative behaviour by outgroup members	(Filak, 2004; Gorham, 2006; A. Maass et al., 1989; Anne Maass, 1999)
Favourability	A general term to describe more favourable coverage to a candidate, party, or ideology	(Eveland & Shah, 2003; S. A. Smith & Roden, 1988)
Front page, Homepage, First story	The biasing of news by journalists to get an article or segment onto the front page, homepage or more air-time	(Baron, 2006; J. R. Zaller, 1999).

9.4 Appendix D – Chapter Two: Bias in the Dissemination of News

9.4.1 Bias in the Dissemination of News on the Radio

Table 9-6 depicts the limited range of biases which have been investigated in the dissemination of news on the radio. Despite an exhaustive search no more studies could be identified in this domain. This is likely due to the issues relating to the comparative difficulties in accessing content from local or even national stations, and huge difficulties to be overcome due to the format of radio news content compared to print content, and the lack of reward versus effort compared to television content, see section 2.3.5.2 later in this chapter for a more detailed description.

Table 9-6 Categorisations of bias that have been investigated in the dissemination of news on radio, their manifestation and effect, including references to relevant papers where found.

Type of Radio Dissemination Bias	Manifestation / Effect	Paper
Presenter Style	Conservative v liberal support. Or combative style on controversial or divisive issues. Supportive or combative interview technique.	(Barker, 2002; H. E. Freeman et al., 1955; McMillan, 2005)
Vocal Characteristics	Use of vocal characteristics such as intonation and tone to denote sarcasm, surprise, incredulity etc. to influence debate or to support or dismiss an idea, policy, or individual.	(Green-Pedersen et al., 2017; Moss, 1988)
Coverage	Includes the duration, and repetition of segments. It also includes the prominence or position in the news cycle	(Green-Pedersen et al., 2017; Larson, 1989)
Structure	The structure of a news story or segment which can intentionally or unintentionally promote or stories or salient points in them	(Berry et al., 1993)

9.4.2 Bias in the Dissemination of News in Print Newspapers

Table 9-7 depicts many of the biases which affect the dissemination of news in print newspapers and news magazines. Bias in print journalism has the longest history of study and as such there is extensive depth to the literature and less overlap and confusion in the categorisation than with television and the WWW. Reasons for the large volume of research into bias in print journalism include the ease of analysis due to the format of the content, comparative ease of access, no requirement to transcribe or annotate, and the single daily snapshot of events compared to online, making longitudinal comparison easier.

Table 9-7 Categorisation of bias investigated in the dissemination of news in print news newspapers and magazines, their manifestation and effect, and references to publications where found.

Type of Print Dissemination Bias	Manifestation / Effect	Publication
Position, Placement, Prominence	Placement of articles in a newspaper or news magazine such as front page, above the fold, or buried deep within the paper. Position or placement bias can also manifest in the placement of facts within a story	(Ahmed et al., 2012; Kenney & Simpson, 1993; Lagun & Lalmas, 2016; Ma et al., 2016)
Layout, Spread	Placement and grouping of articles, photos, headlines to create an initial impression as viewers scan the news or to frame a story before they read it	(Bernstein, 2004; Schindler et al., 2017)
Headline, Lead, Lead Story, Bulletin,	Use of language, tone, focus, respect, endorsement, or condemnation to frame the readers thinking, whether they read the article or not	(Kahn & Kenney, 2002, 2002; S. Kingsbury & Hart, 1933; Kriesberg, 1946; Lott & Hassett, 2014; Merron & Gaddy, 1986; Sachsman, 1970; Stempel III, 1961; Stoodley, 1960; Tannenbaum, 1953; Van Dijk, 2000)
Coverage, Space	The amount of coverage a topic, issue, political party or individual receives. Closely related to position / placement / prominence bias This is also listed as a type of journalist / editorial bias	(Brandenburg, 2005, 2006; D. D'Alessio & Allen, 2000; S. Kingsbury & Hart, 1933; Klein & Maccoby, 1954; D. T. Lowry, 1973; Nokelainen & Kannianen, 2018; Puglisi & Snyder, 2011)
Visual	Using visual methods such as flattering images or caricatures to support or diminish one side in an argument or debate. It may also include the positioning of related news articles and imagery to influence opinion	(Goodnow, 2010; Groeling, 2013a; C. Robinson, 2016)
Photograph	Favourable or unfavourable converge of an individual or event, or the attempt to convey a particular impression, by presenting an individual or event in a certain light or from a certain viewpoint.	(Abraham & Appiah, 2006; Barrett & Barrington, 2005; Greenwood, 2005; Heuer et al., 2011; Kenney & Simpson, 1993; Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Waldman & Devitt, 1998; Zelizer et al., 2002)
Graphics	Biased reporting of information by emphasising certain data while ignoring or diminishing the prominence of relevant counter information to support an argument	(Dick, 2015)
Description	Word choice, use of titles, tone, endorsement or condemnation which may contribute to erroneous reporting of events or a disrespectful representation of an individual, party or policy	(Earl et al., 2004; McCarthy et al., 1996, 1999; McCarthy & McPhail, 1997)
Labelling	Use of titles, nicknames or other terms of disrespect. Highlighting certain individuals background, religious or political beliefs	(Eisinger et al., 2007)
Symbols, Symbolism	This can include the use of iconography or symbols to represent an organisation in a certain manner the continuous use of certain visual representations to depict a certain group, race, culture or religion	(Bruter, 2009)
Quotations	Erroneous, slanted, or out of context quotations	(S. Kingsbury & Hart, 1933)

9.4.3 Bias in the Dissemination of News on Television

Table 9-8 contains the categorisations of bias that have been studied in relation to news on television. Despite best efforts to delineate between them, there are large areas of overlap. One reason for this is that many researchers use interchangeable terms to describe the same bias, e.g. visual or presentation bias. In many instances a single study investigates multiple forms of bias and it is therefore listed under each. For the most part, the studies are categorised as the author described the form of bias it was investigating.

Table 9-8 Categorisations of bias that have been investigated in the dissemination of news on the television, their manifestation and effect, including references to relevant papers where found.

Type of Television Dissemination Bias	Manifestation / Effect	Paper
Coverage within a News Cycle	Includes the amount, time, prominence, duration, repetition, and serial position of segments. It also includes the amount of promotion a segment gets. This can impact the recall of a news item in the viewer	(Groeling, 2008; Gunter, 1979, 1980; Lo et al., 1996; J. E. Newhagen & Reeves, 1992)
Soundbite or Imagebite	The choosing of sound or image clips, often as an advertisement to a larger story, which may frame or set the agenda for a story or make an individual or cause appear in a favourable or unfavourable manner	(Esser, 2008; Grabe & Bucy, 2009; Lo et al., 1996; D. T. Lowry & Shidler, 1998)
Story Format, Debate Style, Discussion Method	The use of different story formats, such as balanced view, or inference story, or discussion formats such as lecture, interview, or discussion, to emphasise different aspects of a news story or to frame it in a certain manner	(Brandon, 1956; Kline, 1981)
Program Format	Differences in the formats of new news shows such as cable news, political satire, and late night news talk shows on the news they produce	(Hoffman, 2013)
Factual Presentation in the Message	The presentation of certain 'facts' about a news story. What constitutes a 'fact' may be subjectively interpreted by the news anchor who may also provide unbalanced coverage	(Doll & Bradley, 1974; Efron, 1971; Russo, 1971; Stevenson et al., 1973)
Editorial or Journalistic Opinion in the Message	The knowing or unknowing insertion or addition of journalists or news editors' personal opinions to slant the news	(Feldman, 2011; Hackett, 1984)
Objectivity in Message	Was the message or content objective in terms of time, neutral voice, objective language, evidence based, and given equal time in the news cycle	(Doll & Bradley, 1974)
Language, Word Choice, Level of Respect, Rhetorical Formats such as: Sarcasm, Hyperbole, Sneer, Sardonicism	The use of biased language to identify or differentiate against a particular group or cast them in a disparaging light	(A. Maass et al., 1989; Anne Maass, 1999; Rainville & McCormick, 1977; Sendén et al., 2014, 2015)
Facial Expressions	Perceived bias in the facial expressions of news anchors which along with other non-verbal cues can communicate scepticism, disdain, disgust, sarcasm, etc.	(H. S. Friedman et al., 1980) (Miller et al., 2007; Mullen et al., 1986; Zimmerman, 2013)
Personal Presentation and Clothing	The personal appearance of the news anchor which has been shown to make a message to be perceived more credible and less biased. It also includes the personal appearance of the subject i.e. prison uniform or three-piece suits	(Barnett, 2003)
Physical Characteristics and attributes	The impact of physical characteristics of the source, typically a news anchor, on the perception of bias. Typically measured as part of credibility	(Slater & Rouner, 1996)
Audience Response Information	Unbalanced or slanted accompanying text and graphics which may be overlaid on the news such as viewer polls	(Davis et al., 2011; Saks et al., 2016; J. B. Weaver et al., 2009)
Presentation	Overarching term to describe news media bias. Often involves the positive or negative representations of an individual, situation, or organisation by a number of visual, linguistic, or other means to create a reality different from the truth	(Groeling, 2013b; Hackett, 1984; Kline, 1979; W.-Y. Lin et al., 2011)

Visual	The visual image presented by a news anchor or news agency, of an individual, candidate, party, or organisation to communicate a favourable or unfavourable representation or frame a story	(S. Banning & Coleman, 2009; Barnett, 2003; Grabe, 1996; Groeling, 2013a; Kepplinger, 1982)
Vocal Characteristics such as: Speech Rate, Intonation, Tone, Pitch, Prosody, Cadence, Accent	The use of paralinguistic elements of speech to influence the content of the message being conveyed to the consumer	(Wortham & Locher, 2009)
Headline, Lead, Lead Story, Bulletin	Use of words, tone, focus, respect, endorsement or condemnation combined with image bites or soundbites to frame the readers thinking whether they read the article or not	(Boyd, 2016; Brandenburg, 2005)
Favourableness or Degree of Polarity +/- of the Content or Message on a Scale	The degree to which statements made in a television news broadcast are favourable / unfavourable to one candidate, party, ideology or position	(Lo et al., 1996; D. T. Lowry, 1973; Turner, 2007)
Nonverbal Communication, Gesturing	Communicating nonverbally in a favourable or unfavourable manner towards a person or situation	(Babad, 1999; Meadors & Murray, 2014; Weisbuch et al., 2009)
Symbols, Symbolism	This can include the use of iconography or symbols to represent an organisation in a certain manner the continuous use of certain visual representations to depict a certain group, race, culture or religion	(Pride & Wamsley, 1972)
Chyrons	Refers to the captions or graphics that occupy the 'lower third' area of a screen during a news broadcast, see Figure 9-3	(N. L. Smith, 2014) Non-Academic ⁴⁹

Figure 9-3 shows two screen shots from CNN and MSNBC with chyrons. Chyrons, also referred to as the 'lower third', have become increasingly important as news editors have realised that televisions displaying twenty-four-hour news are often turned down. Traditionally these have been used to identify the person or scene, set context, or to provide the news headlines. However, recently they have been increasingly used to frame the news that is being displayed to the viewer (MSNBC), or for real time fact checking (CNN). The CNN chyron has become infamous as one of the first examples of real time fact checking, and for its method and language. The MSNBC chyron is an example of framing an event within the context of another.



Figure 9-3 Examples of chyrons from CNN and MSNBC.

⁴⁹ While there is a limited amount of academic research on bias in chyrons, they have recently found a new purpose as real time fact checkers. Consequently, they are facing increased scrutiny in news media. A recent Washington Post news article highlights several issues with their use and how they can frame or bias viewers thinking on a subject. This area will likely receive a lot more academic attention in the near future: https://www.washingtonpost.com/graphics/2018/lifestyle/style/how-cable-news-chyrons-have-adapted-to-the-trump-era/?utm_term=.7ca1297941fd

9.5 *Appendix E - Chapter Two: Studies on Bias in the Dissemination of News in Print Newspapers and News Magazines*

9.5.1 *Coverage Bias*

Kingsbury and Hart undertook one of the oldest series of studies on bias in the dissemination of news in traditional print mediums (S. Kingsbury & Hart, 1933). They analysed the amount of coverage provided to different sides in debates in the US congress in 10 US newspapers. Their first study focused on the debate over a spending bill to fund the construction of 15 cruisers for the US Navy. This was a direct bone of contention between those arguing in favour of preparedness versus internationalism, with the isolationists wanting to stay out of wars in Europe. Their study demonstrated that 8 of the 10 papers provided more space to the argument of congressman Brookhart compared to Congressman Reed. While 6 out of the 8 were within a small margin, two of the newspapers dedicated substantially more space to congressman Reed than Brookhart. 33.3 inches to 3.9, and 44.1 inches to 2.5, a ratio in favour of Reed by 8.8 and 17.6 respectively. While the study goes on to look at the language used and the facts that were omitted, the initial focus is noteworthy as it marks the earliest study investigating coverage in the dissemination of news. A review of the work and criticism of the study can be found in contributions by MacGill Hughes and Murphy respectively (MacGill Hughes, 1938; Murphy, 1934). Coverage bias has received continuing empirical attention, including studies on coverage of newspapers front pages (Klein & Maccoby, 1954; D. T. Lowry, 1973), a meta-analysis of 59 studies on party coverage in US elections in television, newspapers and news magazines (D. D'Alessio & Allen, 2000), study of coverage bias of political parties in newspapers and television of the Irish general election in 2002 (Brandenburg, 2005), and coverage of the main political parties in 7 UK newspapers during the 2005 general election (Brandenburg, 2006).

9.5.2 *Headline Bias*

Studies into bias in news headlines in print media have a long history in the literature. Headlines are often more susceptible to the influences of bias, being that they are designed to attract attention yet are limited in length. Noting these issues led Tannenbaum to state: "*Obviously, bias can enter here, virtually through an open door.*" (Tannenbaum, 1953). They go on to note that many readers are essentially: "*shoppers for headlines*" and highlight the power of journalists who write headlines in shaping public opinion as many readers do not read past them. This they argue, is an even stronger means by which bias can be introduced and affect the consumer. The study found that headlines have a definite effect on the interpretation of the story, especially when the articles are only lightly or superficially read. They also note that this effect may be more pronounced in a real-world environment where there are usually groups of headlines and other supporting material such as in a spread of news articles. Kingsbury and Harts' previously mentioned paper contained a quantitative and qualitative analysis of news headlines. The quantitative analysis was essentially a coverage analysis of the size of each headline. The qualitative analysis involved dividing them into 'fact points' for comparative analysis (S. Kingsbury & Hart, 1933). The study showed that the majority

of newspapers favoured Senator Reed (who advocated preparedness) over Brookhart (who advocated internationalism / unpreparedness) both in coverage and in favourable representation. In a study focusing on news about Soviet Russia in the New York Times (NYT), Kriesberg accounted for the size and position of headlines. He also found that: “*Unwarranted headlines*” were consistently unfavourable to Soviet Russia and argued that: “*there is no attempt to judge headlines by the criterion of ‘true’ news.*” (Kriesberg, 1946). While headlines are essentially textual content, their size and nature mark them much closer to an advertisement for the articles they represent. They are also often heavily styled to fit in with the overall design of the newspaper. As such, they should not be treated strictly as content but should be considered from a design prospective too. Since then a range of studies have looked at bias in news headlines including two studies on the use of favourable / unfavourable ‘loaded’ words (Sachsman, 1970; Stoodley, 1960), a study on headline coverage and ‘play’ (Merron & Gaddy, 1986), size, prominence and favourability of headlines (Kenney & Simpson, 1993), a discourse analysis including the rhetoric of headlines (Van Dijk, 2000), analysis of the tone (Kahn & Kenney, 2002), and positivity and negativity (Lott & Hassett, 2014). Many of these used headline analysis frameworks such as the Stempel classification system (Stempel III, 1961). The most recent contribution, and the most in depth, is that of Naiva and Osorio who studied political bias in newspaper headlines in Chile from 1994 to 2010 during the first 100 days of each government in the two main daily newspapers. They also analysed the headlines for gender bias, bias against each president’s political position, and issue bias (Navia & Osorio, 2015).

Of note to studies on bias in news headlines is Lowrey’s work on the influence of news website design directors on the presentation of news, including news headlines (Lowrey, 2002). In a survey of 230 news website design directors and 224 news editors, they were asked several questions about the presentation of news, including questions relating to the length and placement of news articles, and the choice of accompanying imagery. Both groups of participants were also asked: “*How much control do designers have over what headlines say?*” On a 5-point scale, 1 = low, 5 = high, the Mean score for design directors was 3.15, while the Mean score for news editors was 2.91. While both are the Mean scores from a self-reported survey, it none the less demonstrates that design directors had a not insignificant say in the textual content of newspaper headlines, along with input on the size, placement, and accompanying imagery of news articles. While this study focused on print newspapers, it can be safely assumed that the influence of design directors and designers has only increased in relation to the presentation of news on the WWW due to the increased role and importance of the presentation.

9.5.3 *Photograph Bias*

Photographs have been studied extensively as to how they impact or contribute to bias, especially framing bias. The earliest study relevant to this literature review is that of Moriarty and Garramone who analysed 283 newsmagazine photographs over two months of the 1984 presidential campaign using a 13 point index broken down into categories on Behaviour, Context, and Perspective (Moriarty & Garramone, 1986). They

found that early campaign photos favoured republicans, but the democratic candidate received more favourable treatment in the later stages of the campaign. Since then, a myriad of studies have been undertaken on bias in the photographs in print media, especially relating to presidential campaigns (Barrett & Barrington, 2005; Greenwood, 2005; Kenney & Simpson, 1993; Moriarty & Popovich, 1991; Waldman & Devitt, 1998). There has also been work demonstrating differing use of photographs showing culpability between Palestinians and Israelis in the 2000 Intifada (Zelizer et al., 2002), the stigmatisation of overweight people in images used in online news (Heuer et al., 2011), and the use of visual imagery in the priming of racial stereotypes (Abraham & Appiah, 2006).

9.5.4 *Frontrunner Bias*

Directly relating to bias in the use of photographs and yet only briefly alluded to in several studies, is the concept of frontrunner bias. This was highlighted in Waldman and Davitt's aforementioned study of photograph bias (Waldman & Devitt, 1998). They argue that photographs of candidates more favourably reflected those who were doing well in the polls. The phenomenon was first coined by Johnson in his content analysis of coverage of the 1998 presidential primaries (Johnson, 1993). Baum and Just claim that many scholars in the field use horserace metaphors to describe this phenomenon and note that: "*the impact on voters of the ahead/behind horserace frame has attracted far less scholarly attention.*" (M. Baum & Just, 2009). Most recently, the subject has received some attention from Lin et al. who note that newspapers have more frontrunner slant than blogs, and that news outlets with the strongest frontrunner bias may have a different set of frontrunners than other news services. This fits well with the dichotomy between conservative and liberal media and regional influences (Y.-R. Lin et al., 2011).

9.5.5 *Visual Bias*

The supertype visual bias, see Table 2-2, has received significant attention in print media. It should be noted that many of the aforementioned works on photographic bias also come under the category of visual bias. It is also likely that the term visual bias is receiving increased usage because many of the images used to represent candidates or individuals are not strictly photographs or have been heavily modified⁵⁰. The first relevant contribution is that Goodnow who investigated visual bias in photo essays of Obama and Clinton. Her study noted the differences between colour, saturation, layout, composure and warmth in depictions of the candidates (Goodnow, 2010). In an in-depth study of visual press bias in a multi-party election in New Zealand, Robinson tracked the size and content of images of party leaders including the impact of gender and ethnicity (C. Robinson, 2016). Among many other findings it showed that the female co-leaders of the Green party and the Maori party received far less photographic coverage. It also showed that taken as a whole, coverage of the Maori leaders of parties was less than non-Maori leaders.

⁵⁰ Examples include the "Hope" poster for Barack Obama's US presidential campaign 2008 created by the artist Shepard Fairey, https://en.wikipedia.org/wiki/Barack_Obama_%22Hope%22_poster

9.6 *Appendix F - Chapter Two: Studies on Bias in the Dissemination of News on the Radio*

9.6.1 *Reasons for the Lack of Research into Bias in the Production or Dissemination of News on the Radio*

Research into bias in the production or dissemination of news on the radio has been muted. This is despite many studies on bias in the other mediums highlighting radio as a potential future direction (Latham, 2013; D. T. Lowry, 1986; Merron & Gaddy, 1986). There are several reasons for the comparative lack of research compared to television and especially print. 1) The ease of access, availability, and permanence of content from newspapers. There are also many datasets both academic and private such as the New York Times Annotated Corpus, Factiva, Integrum, and LexisNexis, which can be relatively easily accessed. In comparison, radio content is relatively temporary and difficult to obtain, and corpora are much rarer. 2) There is also an innate difficulty of the spoken word format. News articles are typically written by one or more writers. They are composed, edited and rewritten. In comparison, the dialogue from radio news shows is un-composed and much of the meaning expressed in dialogue, such as intonation and tone, cannot be accurately or easily represented in a written transcript. 3) The format of many radio shows involves multiple parties and multiple points of view, and often shock tactics deliberately designed to elicit response or captivate the attention of listeners. Debate or roundtable discussions can also result in guests and/or the public speaking or even shouting over each other. 4) Groseclose and Milo also claim that newspapers yield more data than other formats such as radio (Groseclose & Milyo, 2005a). 5) Converting spoken dialogue from a radio show into a format usable for many standard types of studies, such as a content or a coverage analysis is also time consuming. At the minimum it would require; recording, editing, transcription, speaker identification, and annotation, all of which are extremely labour-intensive tasks. 6) While many of these issues are also prevalent in television, the visual nature of the medium means that they are easier to overcome. The complexity and range of studies, means the rewards are much greater. 7) There is also a consideration that in the US, where much of the research on bias has been undertaken, many radio stations are relatively local, and it is rare to have a single station or show with national coverage. Syndication of content among different stations is standard. In comparison, the national television networks in the US serve and compete for the same audience making them better suited for comparative experiments. 8) As demonstrated in this thesis, only a minority of studies into bias in the news involve a strict content analysis or the message contained therein. Researchers in many cases prefer to focus on surrounding factors such as accompanying imagery, coverage, or headlines. Consequently, even a full transcription of a news radio show would only be of limited use. Of course, many of these issues are prevalent in television. However, the visual nature of the medium offers significantly more options to conduct research in the domain.

9.6.2 Partisan and Class Bias in Australian Radio

The oldest study identified is that of Aitkin, who investigated partisan bias in Australia's media (Aitkin, 1972). His study found that very few participants found the radio to be biased in favour of any party, but also very few used the medium to actively follow news. Most interesting is the fact that the study was among the earliest to reveal a cultural or class bias in the output. The better educated and more affluent Liberal party supporters were: "*more receptive to the A.B.C.'s consciously 'cultural' broadcasts*", while working class Labour party supporters preferred commercial radio stations whose output presumably was less culturally orientated. While the study does not attempt to identify what aspects of each station broadcasts caused the split in the audience, a class bias was none the less revealed.

9.6.3 Partisan or Pro Government Bias in South African Broadcasting Corporation

Finn undertook a verbal content analysis of radio and television broadcasts before the 1981 election in South Africa. The study found that the South African Broadcasting Corporation (SABC) provided prejudiced and biased coverage in favour of the pro-apartheid National Party (S. M. Finn, 1982).

9.6.4 Identifying Bias in the Speech of Radio Broadcasters

In another Australian contribution, Moss highlights the paucity of research on broadcaster's speech in mass media research (Moss, 1988). His framework is designed to improve the analysis of broadcaster's transcripts to include aspects of discourse such as intonation, combined with their choice of words. His contribution focuses on two types of adjuncts, namely conjunctive adjuncts, which play an important role in narrative structure relating one clause to another in text, and modal adjuncts, which are used to express the speaker's judgement in relation to the message. He believed that such are important to detecting editorial bias and judgements in the text of news bulletins and in interviews and monologues by the radio commentator. His study of radio broadcasts used four main methods of analysis; rhetorical, representational, verbal, and intonation.

9.6.5 Favourability and Coverage Bias in Radio

In an attempt to address perceived issues with studies into media bias, Larson opted for a new approach to assess favourability and coverage bias in radio (Larson, 1989). Her longitudinal study of a single highly acclaimed and award-winning show on National Public Radio (NPR) in the US allowed her to assess bias by determining whether its coverage of presidents, has been more or less favourable than its coverage of other stories in general. Despite the amount of definitions of bias which specifically highlight a time dimension, see Table 9-1 and particularly Entman's definition, see section 9.2.1.1.3, this is the first study of its type encountered in the literature. The study showed that in general, the NPR radio show was negative in its coverage of presidents, and especially the republican presidents Ford and Raeagan. This overt negative coverage may be explained by the fact that news agencies have previously been shown to slant

their news in favour of their audiences (Gentzkow et al., 2015; Gentzkow & Shapiro, 2010; Mullainathan & Shleifer, 2005). Eisinger citing Gentzkow and Shapiro points out that NPR is just as capable of this: *“Because liberals prefer National Public Radio (NPR), the NPR editors and writers are more reluctant to disseminate information that contradicts the liberal point of view.”* (Eisinger et al., 2007). Nixon and Reagan were Republican presidents while Ford was a Democrat president. It must be noted that this was a longitudinal study of a single show on NPR. A more recent study by Groseclose and Milyo comparing bias in news organisations across multiple mediums, found that a different show on NPR called ‘Morning Edition’: *“hardly differs from the average mainstream news outlet”*. This is despite claims by conservatives of NPR being part of the liberal media. It was also considered less biased than the Washington post, NYT, LA Times, and Newsweek among others (Groseclose & Milyo, 2005a).

9.6.6 *Lack of Research on Advertising Bias on Radio*

There has also been a lack of significant work on the impact of advertisers on bias on the dissemination of news on radio. This is despite the fact that, as Shoemaker claims, advertisers have more influence over radio and network television organisations as they receive all their revenue directly from advertisers. *“Television and radio are more sensitive to the need to make a profit than are newspapers and magazines. Virtually all Television and radio income derives from advertisers, and these media compete head to head against similar products in the same market. Virtually every programming decision has economic ramifications.”* (Shoemaker, 1996, p. 257). In comparison, newspapers and news magazine receive at least some of their revenue from the consumer: *“advertising accounts for nearly 100 percent of Television and radio broadcasting, 75 percent of newspaper, and 50 percent of magazine revenues (Dunn & Barban [1986], as reported in Weis & Burke, 1986)”* (Shoemaker & Reese, 1996, p. 211).

9.6.7 *Incumbency Bias on Radio*

In perhaps the largest, most complex, and longest running study relating to radio bias, Green-Pedersen et al. assembled a dataset of more than 30,000 daily radio news features from 1984 to 2003 (Green-Pedersen et al., 2017). Their in-depth study focused on incumbency bias by studying among other things, coverage bias, negativity bias, and tone bias. They argue that outside of elections cycles, news organisations default to a ‘watchdog’ position and focus on social issues and perceived failings of the government. This is done to fight social injustice and highlight inaction or inequality. Consequently, the increased coverage of the main actors or the incumbents is often increasingly negative in tone.

The studies highlighted thus far show that while bias has been looked at in terms of the dissemination of news on radio, the level of attention it has received has been markedly less than the other mediums. Whereas the studies highlighted for television and print represent only a selection of the most relevant contributions to the domain, the studies highlighted here represent all the relevant work on bias on the radio this review could find. This is despite the fact that multiple authors who have not investigated bias on the

radio have highlighted it as a limitation or possible future direction of their work. Also, of particular importance is the fact that talk radio, especially in the US, where much of bias research is conducted, is considered to be predominately conservative. As Lyengar and Hahn citing Pfau et al. state; *“Republicans gravitated to talk radio, radio news, and Television advertising, whereas Democrats avoided talk radio and tuned in to Television news magazines and late-night entertainment Television (Pfau et al., 2007, pp. 36–38).”* (Iyengar & Hahn, 2009; Pfau et al., 2007).

9.7 *Appendix G – Chapter Two: Studies on Bias in the Dissemination of News on Television*

9.7.1 *Presentation Bias in Television News*

Presentation bias is a supertype, see Table 2-2, which includes placement, relative prominence, headlines, and choice of words. It was defined by Lin et al. who highlighted the impact of ‘sensationalism’ or ‘tabloidizing’ of television news due to commercial interests (W.-Y. Lin et al., 2011). They argue: “sensationalism can be viewed as a kind of news bias, slanting news by presenting a particular kind of content, using certain formal features and specific news actors in order to attract more viewers’ attention.” This sensationalism often involves forms of presentation designed to grab the audiences’ attention, the effect of which may be detrimental to learning from or understanding the news. Groeling maintains that: “Presentation bias is the focus of the vast majority of the media bias literature.” (Groeling, 2013a). Though it must be noted that most do not specifically use this term.

9.7.2 *Facial Expression Bias in Television News*

On first instinct, bias in the facial expressions of news anchors does not seem like an obvious focus of what is frankly a large body of work. One reason for this was the substantial increase in television audience figures and the decrease in newspaper sales. The so called, though ultimately unfounded, credibility crisis was also possibly a factor (Cecilie Gaziano, 1988). The earliest study of the type was that by Friedman et al. who looked at perceived bias in the facial expressions of news casters during the 1976 presidential campaign. Participants were required to rate 259 silent clips of news casters, of approximately 2.5 seconds, on a 21-point scale from extremely negative to extremely positive, as they said a candidate’s name (H. S. Friedman et al., 1980). Their study was based on the premise that viewers are perceptive to the broadcasters’ political opinions and that this may be transmitted through their facial expressions as they spoke a candidate’s name. They found that three broadcasters tended to have a more positive expression when talking about Carter than when they talked about Ford. Since then, a series of similar studies have been undertaken on bias in facial expressions on news casters (Miller et al., 2007; Mullen et al., 1986; Zimmerman, 2013). Supported by the aforementioned evidence that news organisations, editors, and journalists compose news to suit their audiences, there is an argument that news casters compose their facial expressions to deliberately influence a message. Swerts and Kraemer found that the intended audience of the news impacted the facial expressions of news readers (Swerts & Kraemer, 2010).

9.7.3 *Soundbite and Imagebite Bias in Television News*

The use of soundbites and imagebites in television news was studied by Esser in four countries and across multiple elections (Esser, 2008). Like headlines in traditional print media, soundbites and imagebites act as advertisements for longer, more in-depth segments of news, and like print headlines they are also prone to bias. Soundbites are often accompanied by an image of the candidate. Esser’s study analysed soundbites

and imagebites from two networks each in the US, UK, Germany, and France, to identify ‘moderately’ or ‘clearly’ negative images of a candidate which they term ‘picture selection bias’. This large study did not find any evidence of imagebite bias in any market, though it noted that German Chancellor, Angela Merkel did receive somewhat negative coverage: “*Contrary to our expectation, we found no evidence for a systematic bad-news tendency on U.S. news. Only the German RTL channel showed a modest preference for disadvantageous image bites, the bulk of which concerned Angela Merkel, who is known for her low telegenic appeal*” Ironically, the authors own bias has come to the fore in this last statement. Imagebites were further studied by Grabe in the context of US elections from 1992 to 2004 showing that republican candidates receive consistently more favourable coverage than democratic candidates (Grabe & Bucy, 2009). Lowry and Shidler also studied bias relating to soundbites in the 1992 and the 1996 campaign (D. T. Lowry & Shidler, 1998), while Lo et al. studied it in Taiwan (Lo et al., 1996).

9.7.4 Visual Bias in Television News

Visual bias is another supertype term, see Table 2-2, or arch term that includes all visual imagery used in communicating news on television. In his paper on the challenges and opportunities in measuring media bias, Groeling highlights many of the key contributions on visual bias (Groeling, 2013a). Due to its visual nature many of these focused on television news, including the works on imagebites by Esser, and Grabe and Bucy (Esser, 2008; Grabe & Bucy, 2009). Many would also consider much of the work highlighted previously on facial expressions as coming under the visual bias super type (H. S. Friedman et al., 1980; Miller et al., 2007; Mullen et al., 1986; Zimmerman, 2013). The first serious contribution, specifically relating to visual bias, was that of Kepplinger. His study of visual biases in the 1976 election television campaign coverage in Germany involved interviewing 151 cameramen from local and national television stations (Kepplinger, 1982). 78% of the cameramen said it: “*is definitely possible*” to make a subject appear in a particularly positive or negative manner, while a further 22% said it: “*might be possible*”. Only 1 out of the 151 cameramen interviewed said it was: “*hardly possible*”. The study goes on to highlight how certain camera angles can contribute to visual expressions of Ease, Power, Clumsiness, Antipathy etc. One of the most important contributions of this work, and not highlighted in any study heretofore, is the cameramen’s ability to capture audience reactions, including; cheering, clapping or other forms of visual approval, or conversely visual forms of disapproval; booing, protest banners, or demonstrations, etc. Grabe used a similar method to analyse coverage of the 1987 and 1989 elections by the South African Broadcasting Corporation. Her analysis of coverage showed that the pro-apartheid National Party received favourable coverage (Grabe, 1996).

Though the importance of visual bias has become increasingly apparent, the difficulty in obtaining data, its measurement, and analysis, compared to studies of other types of bias, or studies of bias in other domains such as print media, means that overall, significantly less work has been undertaken. Groeling also notes this fact and highlights another reason, being the broad availability of content from newspapers compared

to television (Groeling, 2013a). Lowry was also aware of this fact. In his contribution on the validity of his measurement scales, used in several studies to measure content bias in statements made by or about presidential candidates, he briefly highlights a shortcoming in his work in that it made no attempt to measure visual bias (D. T. Lowry, 1986). The paucity in significant contributions on visual bias is evident in the time periods between Kepplinger's work and that of Barnett. Her study focused on the impact of visual biases on the perception of guilt of the accused. Imagery such as orange jumpsuits, handcuffs, or being restrained by police officers resulted in participants evaluating the accused as more threatening, dangerous, and guilty, than those who did not see the same biased visual imagery (Barnett, 2003). Banning and Coleman, in another study relating to US presidential elections, studied non-verbal communication such as facial expressions and appearance, and the construction of a news segment, such as camera angle, focus, distance, and movement. Their study found that fairly balanced coverage was given to each candidate overall (S. Banning & Coleman, 2009).

9.8 Appendix H – Chapter Two: Trust

This section will begin by detailing relevant studies on the impact of website design on perceived trust before highlighting comparable frameworks for understanding how users judge and measure it. The aim of this section is threefold. First, to show that the impact of website design has been looked at among other dimensions and measures of credibility. Second, to highlight any models, theories or frameworks that have been put forward to explain how users judge trust online. Third, to investigate whether any empirical frameworks for measuring trust online exist, and if so, whether they might be of any use to measuring the construct of bias.

9.8.1 Introduction

A general definition of trust as a peer to peer construct from Mayer is: *“the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the truster, irrespective of the ability to monitor or control that other party”* (Mayer et al., 1995). Trust has been defined by Shneiderman in an online business to consumer context in his contribution ‘Designing Trust into Online Experiences’ as: *“The positive expectation a person has for another person or an organization based on past performance and truthful guarantees.”* (Shneiderman, 2000). Further definitions and in depth analysis of the construct as it pertains to WWW applications and ecommerce is provided by (Grandison & Sloman, 2000; Yao-Hua Tan, 2000).

Trust, even more than bias, is considered among the core dimensions of credibility. In fact, Hovland and his communications and change group at Yale, defined Trust and Expertise as the primary attributes of a credible source. They maintain that the credibility given to the arguments of a source is the resulting permutation of these two dimensions. *“In any given case, the weight given a communicator’s assertions by the audience will depend upon both of these factors and the resultant value can be referred to as the ‘credibility’ of the communicator.”* (Hovland et al., 1953, p. 21). As previously stated, credibility is a multi-dimensional construct, however measures such as trust, expertise, believability, accuracy, bias, fairness and completeness of information have regularly been used in studies to measure it, especially online (Bucy, 2003; Flanagin & Metzger, 2007; Johnson & Kaye, 1998).

9.8.2 The Impact of Website Design on Trust

The impact of website design on the perception of trust has received a lot of attention, specifically in the areas of ecommerce and e-health. Trust has long been considered an important factor for websites representing business for which trust was also important, such as banks, legal professions, health care providers, and government organisations. However, it was the advent of ecommerce that brought it to the fore as a design consideration.

Design is considered one of the core dimensions of ‘website quality’ which is used to measure the performance of websites. Website quality is a multidimensional construct, with the number of dimensions varying between three (Rocha, 2012) and as many as twelve (S. Kim & Stoel, 2004a). Often the type of website or its domain impacts the number of dimensions, with websites that sell a product or service typically having more due to the increased interactivity and services they offer, and the resulting increased scrutiny they are under. Dimensions of website quality include; appearance, information quality, service entertainment, fit-to-task, response time, security, interactivness, intuitiveness, and privacy, among others (Barnes & Vidgen, 2000; S. Kim & Stoel, 2004a; H.-F. Lin, 2007; Loiacono et al., 2002). Many of these dimensions are the overarching terms resulting from factor analysis of user interviews or surveys. There are large areas of overlap between them and the exact terminology can change from study to study. Though many use the term design, such as Lin (H.-F. Lin, 2007), it may also be known as visual appeal, appearance, visual design, graphic design, visual interface, or aesthetic treatment. The relationship of design to website quality is analogous to the relationship of bias and trust to credibility. Both website quality and credibility are multidimensional constructs, the dimensions of which vary depending on the situation. It should also be noted that trust has also been identified as a dimension of website quality (Hyejeong Kim & Niehm, 2009; S. Kim & Stoel, 2004a, 2004b; H.-F. Lin, 2007; Loiacono et al., 2002; Wolfenbarger & Gilly, 2003). In fact, the perceived quality of a website is: *“highly correlated with trusting beliefs and intentions”* (McKnight et al., 2002).

Although good website design is synonymous with usability, for brevity, the following brief review only highlights contributions that focus on the design, visual quality, or aesthetics.

A plethora of empirical studies have looked at the impact of website design on trust. A study by Roy et al. was among the first to highlight a strong relationship between interface quality and user trust (Roy et al., 2001). Gefen and Straub investigated the impact of social presence, typified by pictures or personal details of staff, on trust in Business to Consumer (B2C) websites. Their study demonstrated that increased social presence increased trust, which in turn increased purchase intention (Gefen & Straub, 2003). In another study on the impact of dimensions of website quality, including the impact of website design and trust, on customer satisfaction, Lin demonstrated that both were significant determinants of customer satisfaction (H.-F. Lin, 2007). Ganguly et al. also demonstrated that website design positively influences trust, thus increasing purchase intention (Ganguly et al., 2010). In a departure from many in the domain, Ou and Sia maintain that trust and distrust are separate concepts by reason of their distinct cognition (Ou & Sia, 2010). Their contribution includes a framework for testing the impact of specific features of a website’s design on trust and distrust in the user. The study looked at eleven website attributes divided into four categories, however no conclusive findings regarding the impact of specific features of the design were evident. Seckler et al. also investigated the impact of specific characteristics of websites on trust and distrust (Seckler et al., 2015). Their study revealed that ‘graphic design’ was frequently mentioned in experiences

which resulted in distrust, but not in experiences of trust. A similar finding was also found for structure design.

The impact of specific features or characteristics of a website's design on trust have also been investigated. The following paragraphs present a small proportion of the research on logos, colour schemes, and photos.

The impact of website logos on perception of trust has been investigated by Lowry et al. who utilised source credibility theory to design and develop logos which communicate credibility, including expertise and trustworthiness (P. B. Lowry et al., 2014). They also show that the positive impact on perceptions of trust in the user are magnified when the design of the website complements and extends the design. It should be noted that Lowry's work is heavily relied on later in the following credibility section of this thesis. This study also focuses on two other core dimensions of credibility, expertise and trustworthiness.

The impact of colours in website design on perceptions of trust has also been studied extensively. Kim and Moon demonstrated that through the manipulation of design factors, including colour, it was possible to influence perceptions of trust in cyber banking (J. Kim & Moon, 1998). Based on their earlier study, which demonstrated that visual design, including the colour, impacted trust, satisfaction and e-loyalty, Cyr et al. specifically investigated the impact of colour appeal, again on trust, satisfaction and e-loyalty (Cyr, 2008; Cyr et al., 2010). Their eye tracking studies demonstrated that colour is a significant determinant for website trust. In a more in depth study, Pelet and Papadopoulou found that colours that are low in saturation and brightness infuse trust, while colours that vivid and are high in saturation have a negative impact on trust (Pelet & Papadopoulou, 2012).

The impact of photos on trust in ecommerce websites has also been regularly investigated. Steinbrück et al. found that including photos of customer service agents had a significant positive impact on perceptions of trust in online banking websites (Steinbrück et al., 2002). Riegelsberger et al. also found that: "*the trustworthiness of low-trust sites can be boosted by adding a photo*". However, when averaged across websites, they found no significant impact, and if anything found that they had a levelling effect (Riegelsberger et al., 2003). More recently Bente et al. demonstrated that trustworthy photos have a positive impact on potential buyers trust (Bente et al., 2012). Surprisingly, they also found that missing photos and information was worse than untrustworthy photos. Similar results were also found by Ert et al. in their study of photos on trustworthiness in the sharing economy (Ert et al., 2016).

The highlighted papers presented here are but a fraction of the literature on the subject. For a more in depth accounting of the impact of website design dimensions on trust, Karimov et al. present an exemplary synthesis of the literature in the domain (Karimov et al., 2011).

9.8.3 *Theoretical Models and Frameworks of Trust*

The range of theoretical models, theories and frameworks to explain how consumers arrive at decisions of trust online is extensive. This section highlights ten theoretical models of user trust online, only a small proportion in the literature. This should be considered in direct comparison to the dearth of relevant models theorising how users judge bias online. Though it is slightly out of date, the synthesis of the literature on models of adoption by Chang et al. is highly recommended for those interested in the domain (Chang et al., 2005).

- The earliest substantive contribution is that of Tan and Theon (Tan & Thoen, 2000). Their analytical model is based on the idea that transaction trust, is the product of trust in the other party and trust in the control mechanisms of the ecommerce platform. This was based on earlier work with Ganzaroli (Ganzaroli et al., 1999).
- Cheung and Lee's contribution is noteworthy for their presentation of a model and a trust measurement instrument (Cheung & Lee, 2000). They maintain that consumers' trust is affected by two groups of antecedent factors, trustworthiness of WWW vendors and the external environment, and that these are in turn moderated by the consumer's propensity to trust.
- To explain how consumers make decisions of trust online, Egger presents his model of trust for e-commerce (MoTEC) (Egger, 2001, 2000). It describes the design factors that affect consumer's assessment of an ecommerce websites trustworthiness, including pre-purchase knowledge, interface properties, and informational content.
- Mukherjee and Nath used a qualitative modelling framework to develop a model encapsulating the components of trust in online banking. Their study also revealed that opportunistic behaviour has a significant detrimental effect on trust (Mukherjee & Nath, 2003).
- In perhaps the most highly cited contribution, Pavlou integrates trust and risk with the Technology Acceptance Model (TAM) (Pavlou, 2003). In a pair of studies, they demonstrate the robustness of the resulting eCommerce Acceptance Model.
- Corritore et al. provide a review of the domain and propose a model of online trust which includes the perception of credibility, ease of use, and risk (Corritore et al., 2003).
- Koufaris and Hampton-Sosa's model has two main influences, perceptions about the company and perceptions about the website, which is moderated by the users trust propensity (Koufaris & Hampton-Sosa, 2004).
- Bélanger and Carter's model focuses on trust in e-government adoption. It comprises factors such as propensity to trust, trust in the WWW, trust in the government and perceived risk (Bélanger & Carter, 2008).
- As a result of the increased importance of mobile banking, Lin developed a model that accounts for innovation attributes; such as perceived relative advantage, and knowledge based trust; such as perceived competence. He maintains that these form the attitude towards adoption which can trigger behavioural change in intention (H.-F. Lin, 2011).

- Belanche et al. propose an in-depth model to explain trust transfer in the continued use of public e-services. The components of their model are divided up into external cues, such as recommendations, and trust transfer components, such as trust in the public administration (Belanche et al., 2014).

Five frameworks relating to trust have also been proposed by scholars as to how judgements of trust are formed between consumers and businesses online. Combined with the ten models of trust highlighted earlier, it is further evidence of the disparity between the domains.

- Wang and Emurian's contribution begins with an overview of the literature and major concepts in the domain (Y. D. Wang & Emurian, 2005). They present: "*a framework of trust-inducing interface design features articulated from the existing literature*" which foster trust, is divided into four categories; graphic, structure, content and social-cue design. The framework also highlights relevant studies on each feature in the literature.
- Riegelsberger et al. provide an extensive framework for research and design of trust in websites (Riegelsberger et al., 2005). Their work is heavily rooted in existing interdisciplinary research including contributions from the arch domain of credibility, such as that from Fogg, and Fogg and Tsang whose work is covered in detail later in the credibility section of this chapter (Fogg, 2003; Fogg & Tseng, 1999). Their framework has two main attributes, truster and trustee with a range of signals, incentives, and context influences connecting each.
- Although Ou and Sia's contribution is highlighted in the next section as an empirical framework, it is worth mentioning here due to its three-factor framework of trust and distrust determinants, which is similar in form to Wang and Emurian's earlier contribution (Ou & Sia, 2010).
- Karimov et al. extremely detailed contribution presents a framework that classifies trust inducing website features in three broad categories, visual design, social cue design, and content design (Karimov et al., 2011).
- There subsequent study validates the framework for empirical analysis of trust in website design. Li et al. begin by presenting a review of previous contributions in the domain before introducing their holistic framework for trust in online transactions. Their framework focuses on user perception of the institution behind the transaction (F. Li et al., 2011).

9.8.4 *Empirical Tools and Frameworks for Measuring Trust*

Six empirical frameworks, in three broad classifications were discovered for measuring trust online. They can be broken down into measuring the construct itself; measuring the impact of specific features or characteristics of a website that are known to increase or decrease perceived trust; or measuring trust as one dimensions of website quality.

The first approach typically relies on categorisations of statements around the core dimensions of trust, such as; reputation, benevolence, or risk, which are then measured with Likert or Likert type scales.

- The oldest such contribution relevant to this review is a proposal by Cheung and Lee for an instrument for measuring trust online. 41 items were card sorted into 11 categories with high degrees of agreement and construct validity. This resulted in a 36-item scale designed to measure the construct. The tool subsequently performed well in tests for reliability and validity (Cheung & Lee, 2000, 2001).
- In a similar approach, Bhattacharjee presents a 7-item scale to measure trust based on trusties' ability, benevolence, and integrity (Bhattacharjee, 2002).
- Corritore et al. present a tool for measuring trust online based on a model of trust, which includes dimensions of credibility, ease of use, and risk. Initially, 34 Likert scales were used to measure the three dimensions with a confirmatory factor analysis showing that 18 of these were appropriate for use in the final instrument (Corritore et al., 2003).

The second approach adopted by some researchers instead focuses on measuring features and characteristics of a website which could instil trust in the user. These include; usefulness, ease of use, and informational quality.

- Ou and Sia present a theoretical research framework to investigate the antecedents and influences of trust and distrust in ecommerce (Ou & Sia, 2010). They also present an 11-item scale focusing on features of a website's design such as; ease of use, usefulness, and consumer feedback mechanism, which can be used to measure trust and distrust.
- Karimov et al. also adopted this approach, though in a much more in depth manner (Karimov et al., 2011). Their scales allow users to measure; visual design, social cue design, content design, and internal and external e-assurance structures, such as trust seals.

Lastly, online trust may also be measured as one dimensions of website quality.

- One such instrument is WebQual. (Loiacono et al., 2002, 2007). Wolfinbarger and Gilly also measure trust as a separate dimensions of website quality their platform, eTailQ (Wolfinbarger & Gilly, 2003).

9.8.5 Conclusions on Trust

Unlike bias, trust has received significant theoretical and empirical attention both on its own and as a dimension of credibility. This includes significant contributions on the impact of website design on trust and theoretical work underpinning how such judgements are made, and how it should be measured. However, trust is not bias. Even a surface level review of the literature shows the difference in the component dimensions of each, such as risk and benevolence for trust, and selection, coverage, omission etc. for bias. There is also no real overlap in the domains they are measured in. Trust is primarily a concern

of ecommerce and e-health and has thus received the majority of attention within the confines of these domains. Whereas bias is predominantly a concern for news, where it has received almost all its empirical attention. Consequently, the tools and frameworks for measuring trust only serve as a guide to inform the process of measuring bias, thus they cannot be adopted to measure the impact of news website design on perceived bias. It should also be noted that by creating an empirical framework and experiment platform for the design, development, and deployment of experiments to measure credibility, or any of its individual measures such as bias, would have much more utility in the domain.

9.8.6 Summary on Trust

Compared to bias, the impact of website design on perceived trust has received significant attention. This is particularly apparent in the areas of ecommerce and e-health. However, unlike bias, there has been an abundance of models, theories, and frameworks, put forth to explain how users form judgements of trust online. There is also a range of empirical tools and frameworks for measuring trust online, of which seven are highlighted here. They fall into three broad classifications, namely; measuring the construct itself, measuring the impact of specific features or characteristics of a website that are known to increase or decrease perceived trust, or measuring trust as one dimension of website quality.

9.9 Appendix I – Chapter Three: Statistical Power and Effect Size

9.9.1 Statistical Power

Statistical power is the probability of detecting an effect, if the effect ‘is’ actually there, thus avoiding making a Type II error. Statistical Power Analysis tests should be undertaken a priori to estimate the required sample size to detect an effect of a certain size. It should be done a priori as part of the experiment design, or failing that, post-hoc. The results of the power analysis should also be reported with every study.

Undertaking the test a priori will inform the experimenter what sample size they need to detect a specific level of effect. Achieving strong Statistical Power to detect a specific Effect Size can depend greatly on the complexity of a planned experiment. This enables a researcher to reduce or increase the complexity of their experiment based on how many participants they can recruit. Alternatively, they can plan or propose the recruitment of additional participants if necessary.

Undertaking a Statistical Power Analysis test post-hoc will inform the researcher what level of effect they can detect with their given data and sample size. A priori is the better approach as it allows the experimenter to improve or adapt their experiment design to increase its validity if they need to meet a certain threshold. Post-hoc only informs the researcher after the fact, what level of effect they can detect with the data they have. In both instances, Statistical Power is most commonly set to 0.95 with an error probability rate of 0.05. G*Power⁵¹ is a tool designed to compute statistical analysis and Effect Size that has support for a wide range of tests often used on data from experiments measuring credibility (Faul et al., 2007, 2009).

Since 1994 the American Psychological Association (APA) has recommended calculating and reporting Statistical Power and Effect Size (L. Wilkinson, 1999). Yet despite this, many studies in the domain do not appear to report either or both (Chung et al., 2010; K. S. Freeman & Spyridakis, 2004; Schwarz & Morris, 2011). Effect size should be considered from two perspectives, the large and the small. First, Effect Size is the quantitative difference between two groups, therefore, the larger the effect size, the greater the difference between two groups. Simply put, the bigger the effect, the larger the difference. Second, the smaller the effect size, the larger the number of participants necessary to detect it.

To demonstrate this point, if a test has a Statistical Power of >0.95% of being able to detect a small effect such as 0.2, this also translates into having a >0.99% actual power to detect a medium effect of 0.5. In plain English, if the test is strong enough to detect a small effect such as a 0.2 with >95% accuracy, then it has a much higher >99% of being able to detect a medium Effect Size of 0.5.

⁵¹ <http://www.gpower.hhu.de/>

The overarching concept is that Effect Size is a means of quantifying the size of the difference between two groups (Coe, 2002). However, while identifying a large difference in an experiment may be desirable to demonstrate the impact of an intervention, the ability to detect a small effect accurately is also important.

9.9.2 Defining an Appropriate Statistical Power and Effect Size for an Experiment

A persistent question in any domain is defining an appropriate Effect Size for an experiment. In the seminal work on the subject Cohen describes an effect of 0.2 as small, not visible or obvious to the naked eye, 0.5 as medium, or big enough that the difference between two groups is visible or obvious, and 0.8 as large, the difference between two groups is: “*grossly perceptible and therefore large*” (J. Cohen, 1969). Coe provides average effect sizes from various research disciplines, demonstrating the wide variance between domains (Coe, 2002). In his analysis of the problem of defining an appropriate Effect Size in the aligned domain of information retrieval, Nelson concludes that it essentially comes down to deciding how big a difference in precision is necessary, in order to make a real difference (Nelson, 2013).

News websites are often the focus of credibility research online. This is likely due to the importance of credibility to both the purveyors and consumers of news, and the availability of the content to researchers. Given that they receive millions of visits every day, even small changes in their design or visual presentation which impacts perceived credibility are important. Therefore, news organisations would be interested in the ability to detect small effect sizes of circa 0.2 or if possible, even less. However, the law of diminishing returns takes effect. To have a >95% Statistical Power to detect an Effect Size of 0.25 in a large 9x9 experiment, 372 participants are required. To be able to detect a smaller Effect Size of 0.2, the number of participants required increases to 577, a >35% increase in sample size to detect a marginally smaller effect size. Therefore, for research into news websites, the CAFE framework recommends that researchers design experiments, which includes the recruitment of sufficient participants, to detect an Effect Size of at least 0.25 with an error probability rate of 0.05 when conducting research on the credibility of news websites. In domains with less traffic and where small differences would make less of an impact, being able to detect an Effect Size of 0.3 or greater would likely suffice.

Many studies investigating credibility report Statistical Power and/or effect size, but not all. CAFE recommends for best practices that researchers should conduct Statistical Power Analysis tests a priori when planning experiments to see how many participants are required. They should also do post-hoc analysis once they have assessed their data and have excluded any outlier or clearly erroneous submissions. Both desired and achieved effect sizes should be reported. While not always an explicit requirement, top level journals and conferences typically require strong Statistical Power.

9.10 Appendix J – Chapter Three: Crowdsourcing

While some research has shown issues with the accuracy of results from crowdsourcing experiments there have also been significant efforts within the research community to show its benefits (Kittur et al., 2008; Smucker & Jethani, 2011). Previous research into credibility has used methods such as targeted sampling for a specific issue (Johnson & Kaye, 1998), college student populations (Chung et al., 2012), focus groups combined with national telephone surveys (C. Gaziano & McGrath, 1986), and random phone sampling from a single city (Kiousis, 2001b). Participants have been incentivised with cash (C. Gaziano & McGrath, 1986), extra class credit (Corritore et al., 2005), magazine subscriptions (Flanagin & Metzger, 2003), donations to charity (Fogg et al., 2001), or a competitive draw to win \$10, \$25 or even \$50 USD vouchers (K. S. Freeman & Spyridakis, 2004).

Crowdsourcing of participants to conduct experiments online has the potential to significantly broaden the demographic, geographic, and socioeconomic pools with little expense, an issue that has blighted the field in recent times. In their position paper, Henrich et al., review a range of comparative databases used in behavioural science in areas that are all applicable to this research such as visual perception, fairness, categorisation etc. Among its many findings, it highlighted that 96% of the participants in behavioural science research came from countries with only 12% of the world's population, and that it is primarily focused on US college students. Their claim that too much research is reliant upon samples taken from Western, Educated, Industrialised, Rich and Democratic or WEIRD societies is a notable contribution to the field (Henrich et al., 2010b). Research has also shown that not only are the results of online experimentation as accurate as both face to face behavioural testing and volunteers sourced through social media and forums such as Reddit, but that it is also more socioeconomically and ethnically diverse (Behrend et al., 2011; Casler et al., 2013).

More recently, credibility studies have increasingly relied on crowdsourcing participants. Among the earliest was the work of Castillo et al., who investigated Information Credibility on Twitter (Castillo et al., 2013, 2011), Yamamoto and Tanaka, who investigated enhancing credibility judgements in web search results (Yamamoto & Tanaka, 2011), Gupta et al. who demonstrated real time credibility assessment on twitter (Gupta et al., 2014), and Rafalak et al. who investigated the influence of demographical factors on credibility evaluation (Rafalak et al., 2014). Huang et al. also utilised crowdsourcing in their social platform for web credibility evaluation in their attempt to build a large dataset of webpages with multiple credibility ratings (Huang et al., 2013). Some members of the same team also subsequently collaborated to develop a decentralised recommender system for web credibility assessment using collaborative filtering (Papaioannou et al., 2012).

Obtaining a strong Statistical Power while using anything but the simplest experiment design typically requires a far larger number of participants than has traditionally been used in credibility experiments.

However, crowdsourcing offers researchers the ability to access large amounts of participants, from diverse backgrounds, in a relatively short time and for a relatively small investment when compared to methods traditionally employed by credibility researchers. One example of this is the incentive of Fogg et al. who donated \$10 to charity per participant, up to \$5,000, (Fogg et al., 2001). Based on the figures reported, this resulted in circa 605 US participants. Not only was this a significant investment and quite time consuming, but the encouragement by the charity for its members to undertake the experiment would have introduced a significant bias. In comparison, the same investment in crowdsourcing, at the same rate as two implementations of the CAFE framework and platform, would result in just under 3000 participants with much more diverse backgrounds (Spillane et al., 2017b, 2018). Researchers should however be cognisant of the possibility that some of the participants they recruit online may be over exposed to research materials and methods in crowdsourced experiments, and as a result become nonnaïveté (Chandler et al., 2014). This can reduce effect sizes and Statistical Power. Consequently, they may need to pre-screen participants to ensure that they have not partook in significant amounts of crowdsourced research. Most crowdsourcing marketplaces offer this facility. For best practices, the CAFE framework and accompanying platform recommends and supports the recent trend towards crowdsourcing of participants for credibility evaluation experiments, but researchers are encouraged to explore best practices in the domain themselves.

9.11 Appendix K – Chapter Three: Weighted Credibility Score

The majority of research in the domain follows the method enshrined in the literature by Hovland and his team, and expanded upon by subsequent researchers in the domain, that a credibility score is derived by measuring multiple dimensions of the construct, and the resulting value can be referred to as its credibility (Hovland et al., 1953, p. 21). Since then, the vast majority of research in the domain calculates source, medium, and message credibility scores as the Mean score of the various dimensions and measures they have chosen to encapsulate it. As highlighted earlier, there is a litany of issues with this approach, including the choice of, and number of dimensions and measures, and the type of scales used.

There are also variations in how final credibility scores are derived. Most researchers in the domain derive their credibility score as the Mean of the individual measures which reflect the dimensions they believe encapsulate credibility. Common dimensions and measures include Trust, Expertise, Accuracy, Fairness, Bias and Depth of coverage. Each of these are commonly measured in a single Likert, or Semantic Differential Scale etc. Alternatively, some researchers use the Mean scores from multiple scales to measure individual dimensions of credibility (Ayeh et al., 2013; Hong, 2006). The credibility score is then derived as the Mean of these multi scale dimension scores. Both approaches have been demonstrated to be valid and reliable measurement techniques.

However, other researchers have upon occasion unnecessarily, and somewhat erroneously, complicated this process further. One example is Freeman and Spyridakis who, in two studies, examined credibility as a construct with as they claim are its three levels; credibility itself, its secondary components, trustworthiness and expertise, and its tertiary components. For trustworthiness, they include; believable, biased, fair, objective, sensational, and truthful. And for expertise, they claim it includes; accurate, complete, competent, and depth of reporting. Yet it must be noted that their measurement instrument in both studies only measured some of the claimed tertiary measures. It is also noted that these so called tertiary measures were not included in the final analysis for this first experiment but appear to be for the second, where it is not clear whether they contributed to their respective secondary credibility measures scores, or if they contributed directly to the primary credibility measure score (K. S. Freeman & Spyridakis, 2004, 2009). This untested complication of existing measurement techniques for no perceptible gain was unnecessary. It also muddles the domain by making comparisons between studies difficult.

9.11.1.1 An Alternative: The Weighted Credibility Score

The CAFE framework proposes a new approach; weighting individual credibility measures to create the overall credibility score. While this adds complexity to the experiment process the benefits outweigh the additional time, effort, and complexity involved. This is a novel contribution to the domain, as to the best of the authors knowledge, no researchers have proposed deriving a credibility score by weighting individual measures.

As previously stated, the majority of studies in the domain derive their credibility score as the Mean score of the measures which reflect the dimensions. This supposes that each of the chosen dimensions has an equal effect on the user, in every domain, on different types of content and for different dimensions of credibility, source, medium, and measure. This is incorrect for two reasons.

First, the fact that researchers choose different dimensions, which are then realised as measures, to represent the construct of credibility and encapsulate the definition they put forward, demonstrates that in different situations, different measures are more appropriate. The variation, overlap, and patterns of the measures chosen depending on the dimension of credibility, source, medium or message and the type of online resource being investigated, e.g. news, health, or ecommerce, can be clearly seen in the gestalt overview of the domain created by the accompanying repository and categorisation of credibility measures³.

Second, two studies have also shown that participants rely more heavily on different dimensions in different domains. In a study with 2,684 participants to ascertain how users evaluate the credibility of websites, Fogg et al. found remarkable differences in the dimensions of credibility participants relied on when assessing the ten websites from different domains. One example is name recognition and reputation. When assessing credibility in general, 14.1% of participants mentioned this as a core dimension of credibility. However, when assessing ecommerce websites, the proportion rose to 25.9%.

Another example is bias. In general, 11.6% of participants mentioned that bias, particularly information bias, was a specific issue when evaluating the credibility of websites. However, when the results are broken down further, the incidences of users mentioning bias as a core dimension of credibility increased to 23.8% when evaluating opinion/review websites, and to 30.2% when they were evaluating News websites (Fogg et al., 2003). Further demonstrating the fact that bias is important when assessing the credibility of online news, in a recent study users were presented with 25 of the most common measures of credibility in a multiple-choice question. 217 out of 405, or 53.58%, selected bias as a key concern when judging the credibility of news online (Spillane et al., 2018). In contrast, when Rains and Karmikel encapsulated credibility for an experiment on e-health websites they used; believable, trustworthy, accurate, complete, and biased. However, a comparative fit index showed bias was problematic and thus it was discarded (Rains & Karmikel, 2009). This demonstrates that researchers and participants rely on different measures to a greater or lesser extent when judging the credibility of information in different domains. Consequently, the CAFE framework proposes weighting different individual credibility measures.

There are six obvious approaches to weighting measures. 1) The simplest is by the importance of the measure. All the research encountered in the domain appears to support the idea that different dimensions

and the resulting measures are of different importance to the construct of credibility, yet no research was encountered that takes this into account. There is a strong argument that core dimensions and measures such as trustworthiness and expertise should receive a heavier weighting than measures such as currency or usefulness in most situations. 2) As previously demonstrated, different dimensions and measures are relied on to a greater or lesser extent in different domains. Whereas bias is of increased importance in news, it is likely only of limited value in domains such as e-health as demonstrated (Rains & Karmikel, 2009). 3) The dimension of credibility, source, medium, or measure. The reliance on different measures for each of these is evidence that some measures are more important than others. 4) The format or medium of information. Measures such as interactivity and attractiveness are likely more important in experiments with image or video content than they are for other formats. 5) Whether or not the information is likely to be processed by the central route, systematic approach, or controlled processes, or the peripheral route, heuristic strategy, or automatic processes in the ELM, HSM, or C&AHIP (Chaiken, 1980; Petty & Cacioppo, 1986; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Researchers may need to account for motivation, newness and boredom, importance, availability of cues, and the complexity of information or the message contained therein. It is likely that measures, such as depth of coverage, factualness, and ability to explain, are more important than measures such as attractiveness, involvement, or success, depending on which dual-process model of persuasion path through the ELM, HSM, or the C&AHIP is chosen. 6) Lastly, there is striking evidence to show that each person conceptualises the construct of credibility with different core dimensions. Therefore, there is a strong argument that some credibility scores should be derived based on which measures are most important to the individual. One possible means of achieving this is to let experiment participants choose as many or as few measures as they require at the start of an experiment. They could also be asked to rank them in order of importance and their contribution to the final credibility score. While this was virtually impossible or at least extremely difficult with legacy means of conducting experiments, it is easily relatively easily possible with an online experiment platform.

It should be noted again, that to the best of our knowledge, no existing research in the domain weights credibility measures when deriving a final credibility score. It should also be noted that this would require significant testing for validity and reliability. With this in mind, CAFE recommends that researchers consider weighting different measures in order to derive the final credibility score. Done properly, it will be a more accurate measurement of the participant's conceptualisation.

9.12 Appendix L – Chapter Three: Choosing and Reporting Statistical Tests

Most credibility experiments in the literature, whether investigating features or characteristics of a website, comparing mediums, or determining differences between user groups, tend to fall into either the category of Association tests or Group Difference tests. Notable examples include forms of the Chi-Square Test for Association, Kendall's Tau, Spearman's Correlation for Associations tests, and various forms of One, Two-way or Three-way ANOVA, ANCOVA, and MANCOVA for group differences. It is also important that researchers report on whether the data meets the assumptions for each test and if not, why they chose to proceed. Certain statistical test assumptions are considered critical while others are less important, due to the robustness of the test or other factors such as the sample size. In all instances, it is important to report Confidence Intervals and Magnitudes of Effect Size such as Cohen's d, Eta Squared, Partial Eta, or Partial Eta Squared. Following a guide such as Laerd Statistics⁶ can significantly improve statistical analysis and reporting.

Choosing the appropriate statistical test is standard best practice, and most research in the domain adheres to this. However, many studies do not report whether the data met the assumptions for the chosen test. This can include basic requirements such as whether the data was normally distributed, or even the type of data collected. As such, the reader must blindly accept in good faith that the test met all the necessary requirements, and if not, why it was decided to continue with the chosen test. CAFE recommends researchers follow the existing best practice in the domain of fully reporting whether each chosen test met all assumptions. It also recommends that researchers also follow the existing best practice of fully reporting Confidence Intervals and Magnitudes of Effect Size where applicable.

9.13 Appendix M – Chapter Three: Using Visual Aids to Report Individual Measures Making up a Credibility Score

Many studies in the domain provide Mean, standard deviation, variance, and other data for individual measures (Abdulla et al., 2002; Bucy, 2003; Cassidy, 2007; S. M. Choi & Rifon, 2002; Chung et al., 2010, 2012). Thus, it is possible to determine the spread in the rating for each individual measure. It is also possible to ascertain the impact of each measure on the Mean credibility score. This is a comprehensive way and the best practice to report results. Alternatively, some studies only provide the Mean scores for individual measures (Cassidy, 2007). Consequently, users are unable to derive the amount of spread in individual measures. This prevents readers from understanding if the perception of e.g. trust, bias, or fairness etc. was on par with the final credibility Mean score or if one or more measures was propping it up or dragging it down. Even worse, some studies do not provide Mean or standard deviation for any of the individual measures (S. A. Banning & Sweetser, 2007; Eastin, 2006). This prevents the reader from learning almost anything about how the final credibility score was derived.

For each measure of credibility, the Mean and standard deviation should be reported as best practice. This provides the reader with the greatest amount of information about the makeup of the final credibility score. It is also worth noting that not only are measures such as trustworthiness, attractiveness, and professionalism important as dimensions and measures of credibility, they are also important in themselves to website designers and owners in domains such as government, news, e-health, and ecommerce. While the majority of research reports the Mean and standard deviation for each measure, it is not universal.

Researchers should also consider reporting each individual measure using visual aids such as Box Plots or XY Scatter Graphs. These can quickly convey additional information which is not immediately obvious from a table of data. Figure 9-4 depicts an XY Scatter Graph showing 300 participant ratings for six commonly used measures of credibility. Jitter was added to make individual ratings clearer. The Mean and standard deviation for each measure is reported at the bottom and the overall Mean credibility score (43.59) is also shown. The Mean bias score was inverted so that a high bias score contributed to a low credibility score. Compared to even detailed tables of data provided by some studies for individual measures, the XY Scatter Graph shown in Figure 9-4 immediately conveys additional information to the reader in a timely and efficient manner. In this example, the credibility score for this webpage was a relatively modest 43.59. As can be seen in the graph, the majority of participants had a relatively high level of trust in the webpage, but most felt that the information it conveyed lacked in expertise. While its accuracy was high, its fairness was low, and a large proportion of the participants considered it to be relatively or considerably biased. Most also considered the information was not in-depth enough.

This data could be interpreted as meaning that overall, participants thought that the website or organisation behind it could be trusted, but on this article or topic, its expertise was seen to be lacking or low. Because

it performed well on accuracy but badly on fairness, bias, and depth of coverage, it is apparent that while the information that was presented was considered likely to be correct, the reader did not consider that they were fair to the topic overall, and likely did not cover both sides of the argument or debate properly. This resulted in a relatively high bias rating. Combined, these contributed greatly to the relatively low credibility score. It is also obvious that the easiest way to fix this would be to increase the depth of coverage and to pay particular attention to presenting both sides of the argument. This would likely have a positive effect on the overall credibility score.

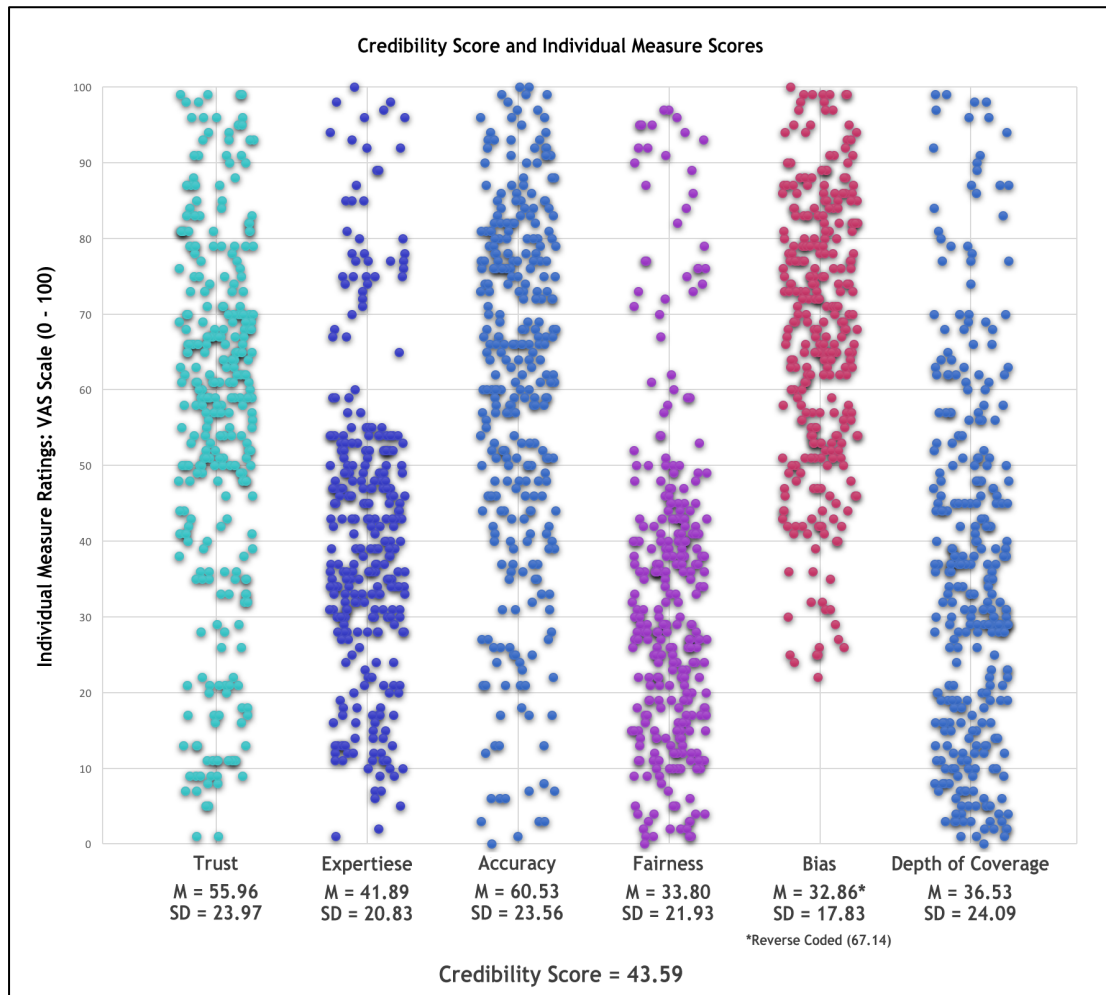


Figure 9-4 Example of an XY Scatter Graph showing the results from an example credibility evaluation experiment with six common measures of credibility which participants (n = 300) rated using VAS scales from 0-100.

By presenting information in this way, readers can make better and informed deductions about the content that was tested and how the final credibility score was derived. The CAFE framework recommends as best practice reporting Mean and standard deviation for individual credibility measures. As a novel contribution to the domain, it also recommends reporting individual measures using graphs or other visual means of communication. It would also be relatively simple to communicate any weighting that was applied to individual measures as recommended in section 9.11.1.1.

9.14 Appendix N - Chapter Four: Experiment One – Statistical Analysis and Results – Initial Bias Rating

Each distorted webpage/article combination was viewed and rated by 15 individual participants. Using a Latin square in reduced form, an incomplete counterbalanced within subject design was achieved. By assigning participants to diagonal paths through the Latin square each participant experienced each webpage/article combination and the effects of each distortion once. Thus, no participant experienced the same content more than once, or experienced the effects of any distortion more than once. In 19 instances, it was not possible to apply the distortion to the webpage/article combination in question due to the absence of the feature in the initial design. In each case these have been clearly labelled with N/A in the corresponding cell in Table 9-9. To conduct the Two-Way Repeated Measures ANOVA, non-applicable distorted webpage/article combinations were replaced with their respective control data. This was not included when conducting the Simple Main Effects for Distortions statistical analysis. Consequently, 62 separate distorted webpage/article combinations form the base dataset.

The following are the results of the participants' initial rating of the positive or negative perceived bias in each distorted webpage/article combination. The re-evaluated bias ratings, which contribute to the final experiment findings, are presented in the section 4.10 later in this chapter.

9.14.1 Main Effects – Initial Bias Rating

A Two-Way Repeated Measures ANOVA was conducted to ascertain the effect of each distortion, on each webpage/article combination, on the perception of bias. An examination of the studentized residuals for values ± 3 standard deviations showed that there were no outliers in the data. A Shapiro-Wilks analysis ($p > .05$) of the studentized residuals showed that the majority of user ratings in the sample were normally distributed with the remainder approximately normally distributed. Mauchly's Test of Sphericity (MTS) indicated that the assumption of sphericity had been violated for the two-way interaction. As the estimated epsilon (ϵ) > 0.75 the more liberal Huynh-Feldt (HF) correction is reported for the two-way interaction rather than the more conservative Greenhouse-Geisser (GG). Multiple paired sample t-tests were not used as they do not have a correction such as the Bonferroni or the Šidák correction which reduce the chance of committing a Type 1 error when conducting multiple comparisons.

The Two-Way Repeated Measures ANOVA revealed a statistically significant two-way interaction between the webpage/article combinations and the distortions, $F(40.14, 562.04) = 1.430$, HF $p = .045$. Thus, Simple Main Effects for Webpage/Article Combinations and Simple Main Effects for Distortions would be investigated.

9.14.2 Simple Main Effects for Webpage/Article Combinations – Initial Bias Rating

To establish if any individual webpage/article combinations impact the perception of bias, Simple Main Effects in the form of a One-Way Repeated Measures ANOVAs were conducted comparing each webpage/article combination to each other at each level of distortion. Where significant results were found, post-hoc pairwise comparisons with the Bonferroni correction for multiple comparisons were conducted. All Confidence Intervals (CI) are 95%. All tests passed MTS except D1.

9.14.2.1 D0

There was no statistically significant difference in the perception of bias when comparing the nine control D0 webpage/article combinations to each other $F(1, 14) = 2.159$, $p = .164$, partial $\eta^2 = .134$.

9.14.2.2 D1

There was a statistically significant effect on the perception of bias when comparing the eight D1 distorted webpage/article combinations $F(1, 14) = 6.093$, $p < .0005$, partial $\eta^2 = .297$. However, it failed MTS $p = .001$, therefore the HF is reported, $p = .631$. As the HF p is greater than 0.05, no examination of the post-hoc pairwise comparisons were undertaken.

9.14.2.3 D2

There was a statistically significant effect when comparing the webpage/articles with the D2 distortion applied $F(1,14) = 5.135$, $p = .001$, partial $\eta^2 = .268$. An examination of the post-hoc pairwise comparisons revealed that:

The Spectator (M -20.73, SE 8.93) was perceived as being significantly more negatively biased than:

- The Economist (M 20.20, SE 8.07), a mean difference of 40.93 (CI 7.46 to 74.40), $p = .012$

9.14.2.4 D3

There was a statistically significant effect when comparing the webpage/articles combinations with the D3 distortion applied $F(1, 14) = 4.367$, $p = .009$, partial $\eta^2 = .238$. An analysis of the post-hoc pairwise comparisons also revealed:

The Spectator (M -16.00, SE 7.91), was significantly more negatively biased than:

- The New Statesman (M 21.80, SE 9.11), a mean difference of 37.80 (CI 8.71 to 66.89), $p = .008$
- The Economist (M 18.80, SE 9.603), a mean difference of 34.80 (CI 1.32 to 68.28), $p = .039$

9.14.2.5 D4

The D4 distortion revealed no statistically significant differences between the perceptions of bias in the five webpage/article combinations tested $F(1, 14) = .478, p = .752, \text{partial } \eta^2 = .033$.

9.14.2.6 D5

There was a statistically significant effect when comparing the webpage/articles combinations with the D5 distortion was applied $F(1, 14) = 6.495, p = .000, \text{partial } \eta^2 = .317$. An analysis of the post-hoc pairwise comparisons also revealed:

The Spectator (M -37.73, SE 7.662) was perceived as significantly more negatively biased than:

- The Telegraph (M 21.47, SE 8.561), a mean difference of 41.80 (CI -100.019 to 16.42), $p = .010$
- The Economist (M 4.53, SE 6.939), a mean difference of 42.27, (CI 7.44 to 77.09), $p = .011$
- New Statesman (M 28.20, SE 9.391), a mean difference of 65.93, (CI 33.85 to 98.02), $p = .000$
- Al Jazeera (M 7.87, SE 7.802), a mean difference of 45.60 (CI 6.07 to 85.13), $p = .016$.
- The BBC (M 21.13, SE 5.91) a mean difference of 58.87 (CI 22.50 to 95.83), $p = .001$.

9.14.2.7 D6

The D6 distortion revealed no statistically significant differences between the perceptions of bias in the six webpage/article combinations tested $F(1, 14) = .478, p = .752, \text{partial } \eta^2 = .033$.

9.14.2.8 D7

There was a statistically significant effect on the perception of bias when comparing the webpage/articles combinations with the D7 distortion applied $F(1, 14) = 3.813, p = .000, \text{partial } \eta^2 = .322$. An analysis of the post-hoc pairwise comparisons also revealed:

The Spectator (M -33.00, SE 9.571) was perceived as significantly more negatively biased than:

- The Economist (M 36.13, SE 8.216), a mean difference of 69.13 (CI 24.61 to 113.65), $p = .001$
- The New Statesman (M 12.47, SE 9.292), a mean difference of 45.47 (CI 9.52 to 81.41), $p = .007$
- Al Jazeera (M 7.47, SE 5.730), a mean difference of 40.47 (CI 2.22 to 78.71), $p = .032$
- Reuters (M 12.60, SE 6.173), a mean difference of 45.60 (CI 1.75 to 89.46), $p = .037$
- Note: The difference between The Spectator and The BBC had a p value of .050

9.14.2.9 D8

There was a statistically significant difference between two webpage/article combinations when the D8 distortion was applied $F(1, 14) = 2.983, p = .005, \text{partial } \eta^2 = .176$. An examination of the post-hoc pairwise comparisons also revealed:

The Spectator (M -13.87, SE 7.665) was perceived as significantly more negatively biased than:

- The New Statesman (M 30.00, SE 7.526), a mean difference of 43.87 (CI 9.44 to 78.30), $p = .006$

Table 9-9 Results of nine One-Way Repeated Measures ANOVAs - Simple Main Effects for Distortions using the Bonferroni adjustment for multiple comparisons.

	D0	D1	D2	D3	D4	D5	D6	D7	D8
Guardian	M -1.20 SE 6.977	M -6.73 SE 5.713 F.273 p .609	M -1.13 SE 8.245 F .000 p .995	M 5.20 SE 7.982 F .267 p .614	M 9.87 SE 5.643 F 1.354 p .264	M 4.07 SE 11.644 F .145 p .709	M 15.13 SE 9.166 F 1.807 p .200	M 12.67 SE 5.836 F 3.156 p .097	M -7.80 SE 6.020 F.466 p.506
Telegraph	M 0.60 SE 5.623	M -7.27 SE 4.615 F 1.175 p .297	M -2.47 SE 4.456 F 2.317 p .150	N/A	M -0.53 SE 7.866 F .011 p .917	M 21.47 SE 8.561 F 4.253 p .058	N/A	M -12.60 SE 7.822 F 1.638 p .221	M 2.80 SE 6.820 F.057 p .814
Independent	M 6.07 SE 8.190	M -16.20 SE 9.959 F 4.501 p 0.52	M -18.73 SE 10.457 F 2.719 p .121	N/A	M 1.53 SE 10.219 F .147 p.707	N/A	M 15.07 SE 9.690 F .686 p .421	M 1.33 SE 7.856 F .139 p .715	M -5.73 SE 7.651 F .897 p .360
Economist	M 12.13 SE 10.154	M 29.46 SE 6.378 F 1.745 p .208	M 20.20 SE 8.065 F .374 p .550	M 18.80 SE 9.603 F 1.187 p .672	N/A	M 4.53 SE 6.939 F .381 p .547	M 6.67 SE 9.886 F .143 p .711	M 36.13 SE 8.216 F 5.532 p .034	M 10.60 SE 8.327 F .015 p.904
Spectator	M -7.53 SE 9.344	M -29.93 SE 13.123 F 1.663 p .218	M -20.73 SE 8.925 F 1.567 p .231	M -16.00 SE 7.913 F 2.294 p.152	N/A	M -37.73 SE 7.662 F 5.946 p .029	M -4.93 SE 11.376 F .025 p .877	M -33.00 SE 9.571 F 3.977 p .066	M -13.87 SE 7.665 F.271 p .610
New Statesman	M 22.27 SE 11.608	M 16.07 SE 9.375 F .199 p .663	N/A	M 21.80 SE 9.117 F .001 p .973	N/A	M 28.20 SE 9.391 F .135 p .719	M 29.60 SE 8.784 F .291 p .598	M 12.47 SE 9.292 F .323 p .579	M 30.00 SE 7.526 F .282 p .604
Al Jazeera	M -0.27 SE 5.483	N/A	N/A	N/A	N/A	M 7.87 SE 7.802 F .897 p .360	M 17.47 SE 8.927 F 2.738 p .120	M 7.47 SE 5.730 F .936 p .350	M 0.20 SE 8.421 F .002 p .967
BBC	M 5.47 SE 7.123	M 18.93 SE 10.001 F 1.158 p .300	N/A	N/A	M 14.67 SE 8.617 F .582 p .458	M 21.13 SE 5.906 F 3.861 p .070	N/A	M 15.40 SE 7.792 F .885 p .371	M 10.33 SE 4.383 F .357 p .560
Reuters	M -7.73 SE 4.318	M -22.80 SE 7.145 F 3.117 p .099	N/A	N/A	M 7.13 SE 10.540 F 1.556 p .233	N/A	N/A	M 12.60 SE 6.173 F 9.045 p .009	M 7.67 SE 9.682 F 1.866 p .193

9.14.3 Simple Main Effects for Distortions – Initial Bias Rating

Simple Main Effects for Distortions were conducted on the participant's initial bias ratings for each distorted webpage/article combination to establish which distortions have a significant effect on the perception of bias. To reduce the likelihood of Type 1 errors while ascertaining the effect of the distortions on each webpage/article combination, One-Way Repeated Measures ANOVAs were conducted comparing

the respective controls of each webpage/article combination (D0) to its distortions (D1-8) using the Bonferroni adjustment for multiple comparisons. Results are presented in the intersecting cells in Table 9-9, with statistically significant results highlighted in grey. The degrees of freedom (DF) for the factor and error were (1,14), respectively. Two distortions, D5 and D7, had a statistically significant effect on three webpage/article combinations.

Only the results of the comparisons between the D0 control webpage/article combination, and their D1 to D8 distorted webpage/article combinations are shown in the Table 9-9, as these are relevant to show the effect of the distortions. Therefore, the comparisons between e.g. D3 and D6 etc. are not reported.

9.14.3.1 D5

The D5 distortion involved the removal of promoted internal content and services from each webpage/article combination. This distortion significantly increased the perception of negative bias in The Spectator from D0 (M -7.53) to D5 (M -37.73). However, it must be noted again that the impact of this distortion is not consistent across all webpage/article combinations for each of the nine websites tested. Promoted internal content is essentially advertisements or clickbait type links provided by third parties such as Outbrain and Taboola. While some of these recommended links highlight relevant serious news articles, many often link to external, less reputable sources. The headlines and image used in such can also be quite prominent and garish, thus influencing participants' overall perception of the news resource.

9.14.3.2 D7

The D7 distortion involved removing article interaction facilities such as social media sharing, email, and voting options. This distortion significantly increased the perception of positive bias in The Economist from its D0 (M 12.13) to D7 (M 36.13) rating, a mean difference of 24.0. There was also a statistically significant increase in the positive perception of bias between Reuters D0 (M -7.73) and its D7 (M 12.60), a mean difference of 20.33. It is possible that the removal of such features signifies to users that the news resource is not reputable and does not invite any interaction with its readers. This is often the practice of websites of extreme left and right perspectives perpetrating to be reputable news resources. However, it must be noted that the effect of the D7 distortion is not common across all the webpage/article combinations tested. The results of these initial assessments of the perceived bias in each of the distorted webpage/article combinations are reported in Table 9-9 above. It shows three statistically significant results: The Spectator: D5, The Economist: D7, and Reuters: D7. These initial results are important for two reasons. First, because they represent the participants' initial perceptions, and second, to compare against the comparative re-evaluated results presented in the next section, and third, as an extra validation step.

9.15 Appendix O - Chapter Four: Experiment One – Statistical Analysis and Results – Comparative Re-Evaluated Bias Ratings – Simple Main Effects for Webpage/Article Combinations

The following are the results of Simple Main Effects for Webpage/Article Combinations conducted on the re-evaluated bias ratings. All Confidence Intervals (CI) are 95%. All tests passed MTS except for D0 and D1, where HF corrections are reported as the estimated epsilon (ϵ) >0.75.

Essentially, these are results of comparisons of the distorted webpage/article combinations between websites. E.g. comparing all nine of the D0 webpage/article distortions shown in Table 4-1 to each other. It should be noted that none of the experiment findings or the experiment hypothesis rely on the results of the Simple Main Effects for Webpage/Article Combinations reported in this section. They are reported here for completeness.

When the results show a statistically significant effect between distorted webpage/article combinations it means that there was a statistically significant difference in the perception of bias between them.

9.15.1.1 D0

There was no impact on the perception of bias when comparing the nine control D0 distorted webpage/article combinations $F(1, 14) = 1.671, p = .113, \text{partial } \eta^2 = .107$. However, it failed MTS $p = .036$, therefore the HF is reported, $p = .779$. As the HF p is greater than 0.05, no examination of the post-hoc pairwise comparisons were undertaken.

9.15.1.2 D1

There was a statistically significant effect on the perception of bias when comparing D1 distorted webpage/article combinations $F(1, 14) = 7.374, p < .005, \text{partial } \eta^2 = .345$. However, it failed MTS $p = .022$, therefore the HF is reported, $p = .672$. As the HF p is greater than 0.05, no examination of the post-hoc pairwise comparisons were undertaken.

The Economist (M 30.67, SE 6.563) due to its high positive Mean bias rating was significantly higher than:

- The Guardian (M -8.33, SE 6.51), a mean difference of 39.0 (CI 11.50 to 66.51), $p = .002$.
- The Telegraph (M -5.933, SE 4.173) a mean difference of 36.60, (CI 4.31 to 68.90), $p = .018$.
- The Independent (M -16.467, SE 9.937), a mean difference of 47.13, (CI 3.62 to 90.64), $p = .027$.
- The Spectator (M -29.93, SE 13.123), a mean difference of 60.60, (CI 11.88 to 109.32) $p = .008$.
- Reuters (M -21.133, SE 7.204), a mean difference of 51.80, (CI 20.20 to 82.64), $p = .000$.

The BBC (M 21.20, SE 10.105) was also significantly higher than:

- The Spectator (M -29.93, SE 13.123) a mean difference of 51.13, (CI 0.72 to 101.54), $p = .045$.

9.15.1.3 D2

The D2 distortion revealed no statistically significant effect between the perceptions of bias in the five webpage/article combinations tested $F(1, 14) = 3.915$, $p = .07$, partial $\eta^2 = .219$.

9.15.1.4 D3

There was a statistically significant effect between the webpage/article combinations tested under D3 $F(1, 14) = 4.583$, $p = .008$, partial $\eta^2 = .245$.

The Spectator (M -16.267, SE 7.879) was perceived as significantly more negatively biased than:

- The Economist (M 20.133, SE 8.065), a mean difference of 36.27 (CI 6.75 to 65.79) $p = .012$
- The New Statesman (M 21.60, SE 9.22), a mean difference of 37.87 (CI 4.15 to 71.58) $p = .024$

9.15.1.5 D4

The D4 distortion revealed no statistically significant differences between the perceptions of bias in the distorted webpage/article combinations tested $F(1, 14) = .618$, $p = .651$, partial $\eta^2 = .042$.

9.15.1.6 D5

There were statistically significant differences between the distorted webpage/article combinations tested under D5 $F(1, 14) = 6.378$, $p = .000$, partial $\eta^2 = .859$.

The Spectator (M -37.667, SE 7.678) was perceived as more negatively biased than:

- The Telegraph (M 20.20, SE 7.56) a mean difference of 57.87, (CI 15.42 to 100.32), $p = .004$
- The Economist (M 5.867, SE 6.83) a mean difference of 43.53, (CI 6.06 to 81.01), $p = .015$
- New Statesman (M 24.60, SE 8.448), a mean difference of 62.27 (CI 13.91 to 110.62), $p = .006$
- The BBC (M 21.00, SE 8.62), a mean difference of 58.67, (CI 25.45 to 91.86), $p = .00$

9.15.1.7 D6

The D6 distortion revealed no statistically significant differences between the perceptions of bias in the six webpage/article combinations tested $F(1, 14) = 1.285$, $p = .280$, partial $\eta^2 = .084$.

9.15.1.8 D7

There was a statistically significant effect on the perception of bias when comparing the D7 distorted webpage/article combinations. $F(1, 14) = 6.278$, $p = .000$, partial $\eta^2 = .310$.

D7 once again confirmed the significant differences between the negative perception of bias in the Spectator (M -33.00, SE 9.571) and three of the other webpage/article combinations with D7 applied including:

- The Guardian (M 11.00, SE 6.59), a mean difference of 44.00 (CI 3.67 to 84.35), $p = .025$
- The Economist (M 36.13, SE 8.216), a mean difference of 69.13 (CI 27.95 to 110.32), $p = .000$
- The New Statesman (M 12.87, SE 9.210), a mean difference of 45.87 (CI 3.17 to 88.56), $p = .028$

The Economist (M 36.13, SE 8.216) was also seen as being significantly more positively biased than:

- The Telegraph (M -13.00, SE 7.764), a mean difference of 49.13 (CI 4.79 to 93.48), $p = .022$.

9.15.1.9 D8

There was a statistically significant difference between two webpage/article combinations when the D8 distortion was applied. $F(1, 14) = 3.047$, $p = .004$, partial $\eta^2 = .179$. An examination of the post-hoc pairwise comparisons showed that:

The New Statesman (M 29.80.00, SE 7.633), was considered significantly more positively biased than:

- The Spectator (M -13.87, SE 7.665) a mean difference of 43.67 (CI 5.17 to 82.17), $p = .018$
- The Guardian (M -9.200, SE 6.515), a mean difference of 39.00, (CI 6.820 to 71.18), $p = .010$

9.16 Appendix P - Chapter Four: Experiment One – Frequency of Re-Evaluation

After individually rating perceived bias in three distorted webpage/article combinations presented in random order from their assigned path through the Latin square, participants were given the opportunity to comparatively re-evaluate and adjust their bias ratings, while viewing the distorted webpage/article combinations they had just rated side-by-side. Of a total of 930 individual data observations, users amended their ratings just 62 times, 34 times to increase the perceived bias and 28 to decrease it. In just three instances, the Mean perceived bias rating for the distorted webpage/article combinations was flipped from positive to negative or vice versa. One of the main reasons behind including the comparative re-evaluation step was to increase the validity of the results by ensuring the participants were happy with their initial bias rating. The low incidences of re-evaluation demonstrate users were confident or happy with their initial bias ratings.

One of the other main suppositions behind the comparative re-evaluation was to discover if there were any trends or patterns behind which webpage/article combinations or distortions were most or least commonly re-evaluated and by how much. However, no patterns or obvious trends were discovered behind the webpage/article combinations. Each was as equally likely to be re-evaluated as the next. The frequency of re-evaluations for webpage/article combinations from each of the nine websites tested are presented in Table 9-10.

Table 9-10 Frequency of re-evaluation of webpage/article combinations from the nine websites.

Webpage/article	Reassessment Frequency	As % of 62	As % of 135	As % of 930
Guardian	6	9.68	4.44	.65
Telegraph	7	11.29	5.19	.75
Independent	8	12.9	5.93	.86
Economist	9	14.52	6.67	.97
Spectator	4	6.45	2.96	.43
New Statesman	10	16.13	7.41	1.08
Al Jazeera	5	8.06	3.7	.54
BBC	5	8.06	3.7	.54
Reuters	8	12.9	5.93	.86
Total	62	100	45.93	6.68

The frequency of re-evaluation for webpage/article combinations are presented in Table 9-11. One trend was apparent when comparing the likelihood of distortions impacting whether users would re-evaluate their perceived bias ratings. The undistorted control, or D0 webpage/article combinations, were far less likely to be re-evaluated than any others. It is possible that participants considered these original versions of the webpages, with the news articles inserted into them as if they were original, as being more balanced, while those with a distortion applied as being more likely to be biased. However, this pattern is weak and would require further confirmation.

Table 9-11 Frequency of re-evaluation of distortions.

Distortion	Reassessment Frequency	As a % of 62	As % of 135	As % of 930
D0	2	3.23	1.48	.22
D1	9	14.52	6.67	.97
D2	5	8.06	3.7	.54
D3	9	14.52	6.67	.97
D4	9	14.52	6.67	.97
D5	9	14.52	6.67	.97
D6	6	9.68	4.44	.65
D7	6	9.68	4.44	.65
D8	7	11.29	5.19	.75
Total	62	100	45.93	6.69

9.17 Appendix Q – Chapter Four: Impact of Socioeconomic Status on Perceived Bias

9.17.1.1 Education

Educational achievement among the participants was very high, with just 10.4% reporting having completed only second level education. 83.7% of the participants completed a minimum of a four-year college degree. In comparison, only 34% of the wider US population have the same (Walther et al., 2004). Education level did not provide any discernible trends as to how participants perceive bias. A two-tailed Z-Test, ± 1.960 SD, showed no significant results, see Table 9-12.

Table 9-12 Education level and perception of bias.

Education Level	% of 930	Mean Bias Rating	SD
High School	10.4	7.51	38.11
Some College	28.5	6.15	37.96
Two Years of College	6.7	0.89	38.02
Four Years of College	38.3	3.51	30.53
Some Graduate Study	6.7	7.17	39.12
MSc or Professional Degree	7.1	-2.89	35.69
Advanced Graduate or PhD	2.4	4.09	21.07

9.17.1.2 Occupation

Occupation results point towards a pattern of young students and early career professionals from upper middle class backgrounds. Only 9.7% of responders reported undertaking skilled, semi-skilled or unskilled manual labour roles. Professional, semi-professional and skilled professional accounted for 54.1% of responders. A high rate, 19.3%, of unemployed individuals took part. Participant's occupation also did not impact significantly on the perception of bias though those with professional type jobs have a more positive perception of bias, a Two-tailed Z-Test, ± 1.960 , showed no significant results, see Table 9-13.

Table 9-13 Occupation and perception of bias.

Occupation	%	Mean Bias Rating	SD
Unskilled Manual	1.5	1.14	35.36
Semi-Skilled	2.8	1.00	37.51
Skilled-Manual	4.6	-1.91	31.05
Skilled-Professional	10.5	4.43	33.76
Semi-Professional	17.1	5.94	36.68
Professional	26.6	3.89	32.81
Unemployed	19.2	6.45	31.55
Stay at Home Parent	0.9	-2.63	32.00
Other	16.8	3.68	41.27

9.17.1.3 Living Arrangements

Living arrangements also point to a student or young professional profile. 20.7% rent in multiple occupancy, 5.2% in a dormitory and 26.7% live with their family. 14.1% and 12.6% reported owning a home with and without mortgages respectively. Although the average user profile is young, it is

representative of the core audience of many news websites. As shown in section 9.19.1.2 and in Figure 9-7 later in this section, 99.3% of the sample report that they access news online. Living arrangements also did not impact significantly on the perception of bias a Two-tailed Z-Test, ± 1.960 , showed no significant results.

9.18 Appendix R – Chapter Four: Experiment One – Elections, Political Leaning and Ideological Viewpoints

Participants were also asked to indicate their political ideology and party affiliation during the initial survey stage of the experiment, Figure 9-5 and Figure 9-6 show a clear liberal and independent trend among participants' ideological values. A large proportion also have an affinity with US Democratic party or claim to be Independents than those with an affinity with the US Republican party.

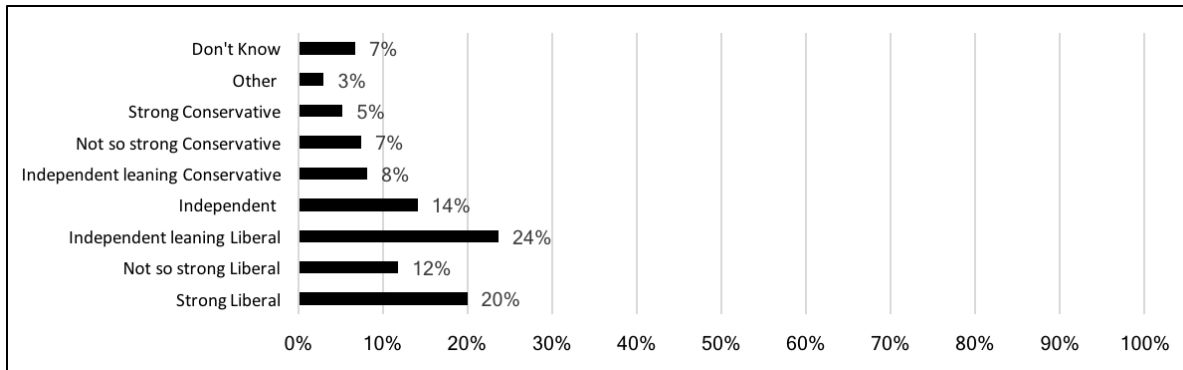


Figure 9-5 Political ideology of participants.

76.3% of respondents believe that voting is important with 33.3% saying it is extremely important. Only 3.0% believe that voting is not important. Figure 9-6 highlights which elections participants regularly vote in.

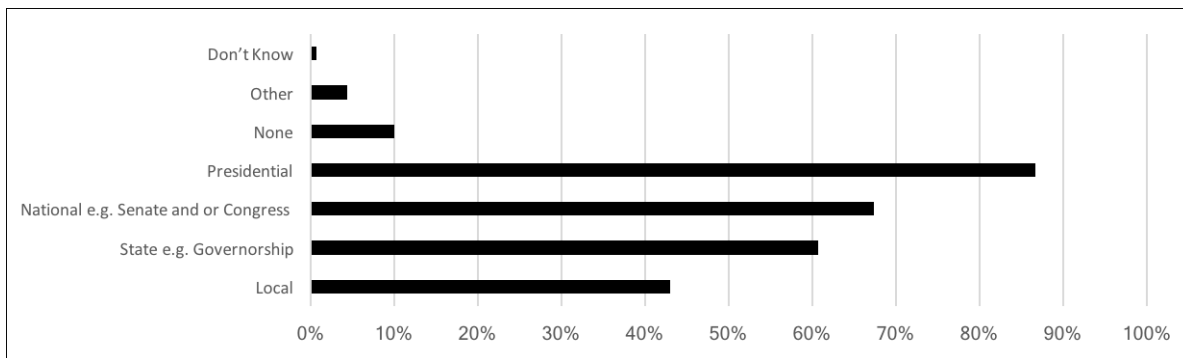


Figure 9-6 Election participations of participants.

Compared to actual rates of participation in US elections, the 2012 Presidential participation rate was just 53.58% of the voting age population, the responders to this experiment claim a participation rate of 86.7%, which would rank them among the highest rates of participation in the world (Pew, 2015b). Whereas 67.4% of participants claim that they vote or intend to vote in National elections for the Senate and Congress, the U.S. census reports that actual participant rates for Congressional elections ranges from 61.8% in 2012 to

41.9% in 2014 (Bureau, 2016). It should be noted that participation rates for Congressional elections are much stronger in presidential election years (File, 2015). Based on these figures, the experiment sample is much politically active than the normal population or at least they claim to be.

9.18.1.1 Political Ideology

A surprising finding in the research is the fact that those with stronger political views, both liberal and conservative, perceive a stronger liberal bias than their less partisan peers, see Table 9-14. When the results of the two groups are compared, they are almost identical, Liberals (all) N = 55.5%, M = 3.76, SD = 34.04 and Conservatives (all) N = 21.3%, M = 3.77, SD = 34.34.

Table 9-14 Political ideology and the perception of bias.

Political Ideology	%	Mean Bias Rating	SD
Strong Liberal	20.3	6.09	35.09
Not so Strong Liberal	11.6	4.23	35.49
Independent Leaning Liberal	23.5	1.30	32.38
Independents	14.0	0.18	28.90
Independent leaning conservative	8.3	4.19	34.16
Not so Strong Conservative	7.5	0.94	35.18
Strong Conservative	5.5	7.02	33.78
Other	2.8	15.15	37.97
Don't Know	6.5	15.82	49.69

9.19 Appendix S – Chapter Four: Experiment One – WWW and News Consumption Habits

9.19.1.1 WWW Usage

WWW usage among the respondents was extremely high. 97.8% report browsing the WWW for a minimum of two hours per day with 34.1% saying that they use it for more than eight hours each day. Of course, the participants for this survey were recruited online, but the figures reported are only slightly higher than Pew’s longitudinal Internet access survey, which report 96% of 18 - 29 year olds, using the Internet each day in 2015 (Pew, 2015c). The majority, 94.1%, also own a smartphone with 78.5% saying that they browse the Internet for a minimum of one to two hours per day on it. This is higher than the two thirds adoption rate for the population as a whole but this is likely due to the younger age profile and economic demographic (Pew, 2015a).

9.19.1.2 News Mediums

Participants were asked to select via multiple choice, which mediums they use to access news, see Figure 9-7. Online was the clear favourite with 99.3% selecting it, while just 17.6% selected print newspapers. Although results were likely skewed due to the fact that participants were sourced from a crowdsourcing marketplace, the extremely low print news readership demonstrates the continuing decline of the industry.

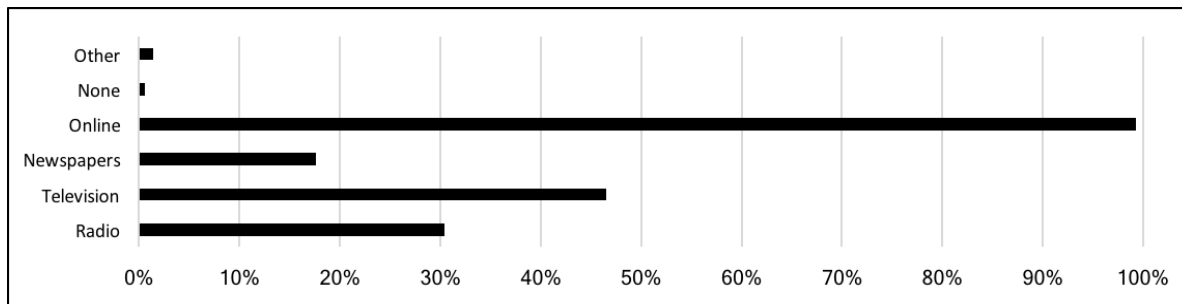


Figure 9-7 Most common mediums participants use to access news.

A Two-tailed Z-Test, ± 1.960 , showed participants choice of news medium did not significantly impact their perceived bias, see Table 9-15. This question was asked because it was thought that perhaps participants who rely on different mediums might have different standards or perceptions of bias.

Table 9-15 News medium access habits and the perception of bias.

News Mediums	%	Mean Bias Rating	SD
Radio	30.4	3.90	34.69
Television	46.5	6.74	38.00
Newspapers	17.6	3.95	34.91
Online	99.3	4.36	35.01
None	0.6	-3.33	7.01
Other	1.4	6.31	28.80

9.19.1.2.1 *Television News*

Participants were also asked, via multiple-choice, which television stations they regularly access, see Figure 9-8. This was asked in order to establish if there was a correlation between participants who access conservative or liberal stations and their perception of bias. There were no significant results comparing television station preferences and perceptions of bias.

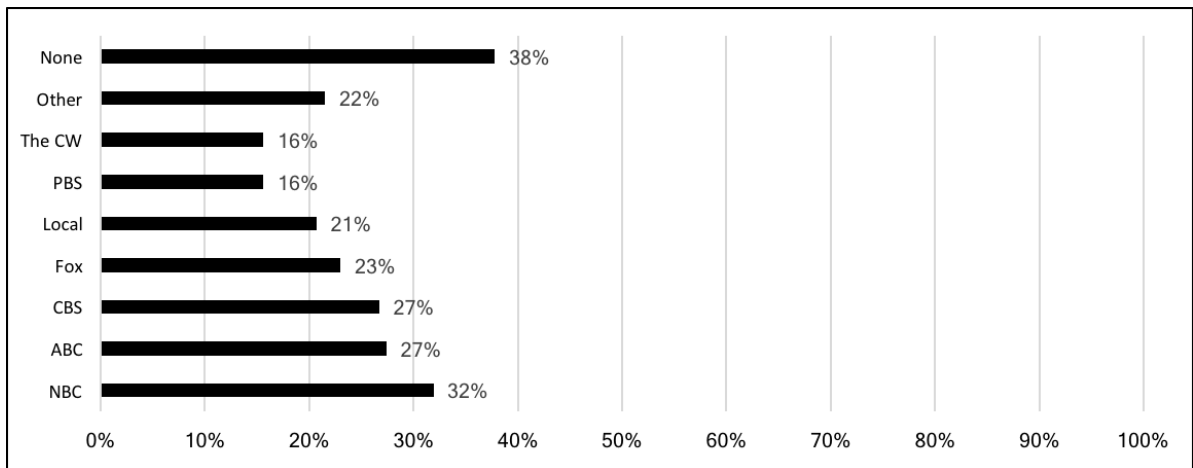


Figure 9-8 US Television news access habits of participants.

9.19.1.2.2 *Newspapers*

Participants were also asked via multiple-choice which print newspapers they read most often for news, see Figure 9-9. This was asked to determine whether there was a correlation between participants who regularly read conservative or liberal newspapers, and their perception of bias. As can be seen in Figure 9-9, over 52% do not access any print newspapers. This is further evidence of the trend away from print media towards other mediums, particularly online which was seen previously in Figure 9-7.

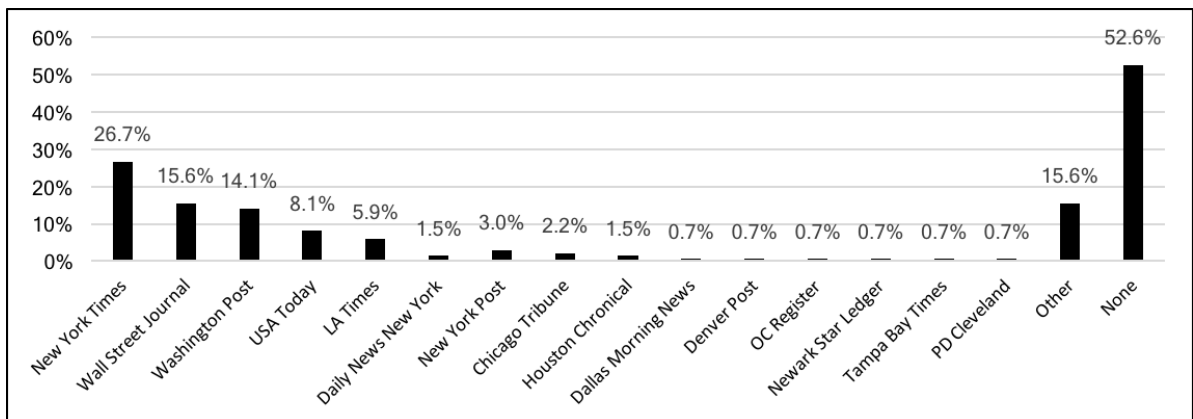


Figure 9-9 Newspaper readership of participants.

Table 9-16 Print newspapers and the perception of bias.

Print News Papers	%	Mean Bias Rating	SD
Wall Street Journal	152	0.68	39.17
New York Times	250	4.32	38.34
USA Today	79	14.61	42.53
LA Times	56	6.29	31.91
Washington Post	136	1.71	35.24

Several methods have been used to determine the Conservative / Liberal position of television news agencies and newspapers. One is that of Groseclose and Milyo who compute scores based on how often each publication cites liberal or conservative think tanks and policy institutes and then compares it with how often conservative and liberal members of congress cite the same groups (Groseclose & Milyo, 2005d). Latham extended this method to also include whether or not the news agencies included warning about any affiliation to the think tanks or their independence (Latham, 2013). While Krestal demonstrated a successful method to compare news articles with political speeches to ascertain their conservative / liberal position (Krestel et al., 2012).

There were no correlations between which newspapers that participants read regularly and their perception of bias, see Table 9-16. Only newspapers which were selected by at least 50 participants were included in the analysis.

9.20 Appendix T – Chapter Four: Experiment One – Reflective Questions

9.20.1 Least Biased Medium

Table 9-17 Least biased news medium.

Question: Please select which news medium do you believe is the least biased?		
Input: Radio button		
Medium	Responses	% of 135
Newspaper	54	40%
WWW	53	39.26%
Radio	19	14.07%
Television	9	6.66%

9.20.2 Awareness of Certain Types of Bias

Table 9-18 Awareness of gender bias in online news.

Question: To what extent are you aware of gender bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	20	14.81%
Sometimes	52	38.52%
Rarely	51	37.77%
Never	12	8.88%

Table 9-19 Awareness of political bias in online news.

Question: To what extent are you aware of political Bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	59	43.7%
Sometimes	61	45.19%
Rarely	13	9.63%
Never	2	1.48%

Table 9-20 Awareness of racial bias in online news.

Question: To what extent are you aware of racial bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	28	20.74%
Sometimes	65	48.15%
Rarely	31	22.96%
Never	11	8.15%

Table 9-21 Awareness of climate change bias in online news.

Question: To what extent are you aware of climate change bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	34	25.19%
Sometimes	66	48.89%
Rarely	28	20.74%
Never	7	5.19%

Table 9-22 Awareness of health care bias in online news.

Question: To what extent are you aware of health care bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	18	13.33%
Sometimes	58	42.96%
Rarely	45	33.33%
Never	14	10.37%

Table 9-23 Awareness of religious bias in online news.

Question: To what extent are you aware of religious bias in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	42	31.11%
Sometimes	58	42.96%
Rarely	25	18.52%
Never	10	7.41%

Table 9-24 Awareness of political party favouritism bias in online news.

Question: Does one or another political party benefit more than the others due to biased coverage in online news?		
Input: Radio button		
Answer	Responses	% of 135
Always	36	26.67%
Sometimes	17	12.59%
Rarely	11	8.15%
Never	71	52.59%

9.20.3 Political Topics that Regularly Receive Biased Coverage

Table 9-25 Political topics that regularly receive biased coverage online.

Question: Which political topics do you believe regularly receive biased coverage in online news? Please select all that apply:		
Input: Multiple Choice		
Answer	Responses	% of 135
Elections	118	87.41%
Local Issues	46	34.07%
National Issues	122	90.37%
State Issues	70	12.59%
Other	6	4.44%

9.20.4 Features and Characteristics of a News Website and News Article that Convey Bias

These are distinct from the message features such as argument quality, balance etc.

Table 9-26 The features of a news article which convey bias.

Question: Which of the following features of a news article have you detected bias in? Select all that apply.		
Input: Multiple choice		
Answer	Responses	% of 135
News Article (Word Choice, Content, or language)	103	76.30%
News Article Headlines (Word Choice, Content, or language)	100	74.07%
Article Photographs	91	67.41%
News Article Videos	74	54.81%
News Article Headline (Space, coverage, size, position, or prominence)	72	53.33%
News Article Cartoons	69	51.11%
News Article Size (Length or coverage)	68	50.37%
Author Profile	66	48.88%
Article Supporting Content	56	41.48%
Article Related Content	51	37.78%
Image or Photograph Captions	40	29.63%
Article Further Reading Links	40	29.63%
Article Inline Links	38	28.15%
Article Lead or Teaser	30	22.22%
Article Graphics	28	20.74%
Article By-line	21	15.55%
Article Metadata	17	12.60%

Table 9-27 The features of a news websites homepage which can convey bias.

Question: Which of the following features of a news website’s homepage have you detected bias in? Select all that apply		
Input: Multiple Choice		
Answer	Responses	% of 135
Homepage Headlines, Leads, and Links (Word choice, language, word difficulty, grammar)	126	93.33%
Homepage (Spread, position, or grouping of related content, articles or headlines on a single topic or storyline)	116	85.93%
Homepage (Overall combination of news stories, headlines and images)	111	82.22%
Homepage Photographs	104	77.04%
Homepage Headlines, Leads, and Links (Fonts, colour, loudness, or punctuation)	103	76.30%
Homepage Cartoons	92	68.15%
Homepage Headlines, Leads, and Links (Space, coverage, size, position, or prominence)	91	67.41%
Homepage Videos	90	66.66%
Homepage Editorial Headlines, Leads, and Links	85	62.96%
Homepage content type (Celebrity photos or skin pictures)	71	52.59%
Homepage Graphics	68	50.37%
Homepage (Symbols and Graphics)	68	50.37%
Homepage picture essays	62	45.93%
Homepage content type (Serious news, tabloid news, clickbait, celebrity news)	60	44.44%
Homepage Banner, Branding, and Logo	48	35.55%
Captions	41	30.37%

Table 9-28 The characteristics of a news website’s design or a news webpage’s aesthetic that impact perceived bias.

Question: Do the following characteristics of a news website’s design or a news webpage’s aesthetic reduce, increase, or have no effect on your perception of bias? Input: Radio Button			
Answer	Reduce	Increase	No Effect
Seriousness	120	4	11
Professionalism	119	9	7
High Quality Aesthetics	105	13	17
Good Design	102	20	13
Modern	76	18	41
Complementary Colour Scheme	75	4	56
Trustworthy	72	13	50
Good Alignment	72	20	43
Technical Competence or Affordances	71	21	43
Dynamic	68	32	35
Expensiveness	65	49	21
Stylish	54	42	39
Simplicity	50	53	32
Cleanness	50	5	80
Calmness	47	11	77
Bad Design	40	78	17
Active	40	51	44
Good Quality Photographs or Images	40	7	88
Traditional	38	36	61
Conservative	37	35	63
Busyness	36	70	29
Clashing	34	59	42
Interactive	34	22	79
Amateurism	34	74	27
Classical	31	9	95
Bad Layout	30	54	51
Cheapness	28	95	12
Loudness	27	71	37
Cutting Edge	24	26	85
Boldness	23	62	50
Clashing Colour Scheme	21	42	72
Minimalism	20	24	91
Low Quality Aesthetics	19	88	28
Openness	14	16	105
Spacious	14	20	101
Bad Quality Photographs or Images	11	24	100
Gaudiness	9	116	10
Tabloid	8	122	5

9.20.5 *Impact of News Website Content or Focus on the Perception of Bias*

Table 9-29 The impact of news website content or focus on the perception of bias.

Question: Do any of the following types of content or the focus of the news website reduce, increase, or have no effect on your perception of bias in the articles it contains?			
Input: Radio Button			
Answer	Reduce	Increase	No Effect
Focus on Trending or Viral Content	34	61	40
Clickbait Style Articles	11	85	39
Content Farm	34	79	22
Focus on Celebrity News	30	81	24
Sexiness of Pictures	40	75	20
Violence in Pictures	19	71	45
Invasiveness of Pictures	10	86	39
Bare Skin in Pictures	19	75	41
Affordance of Underlying Technology	32	21	82
Focus on Sex	41	51	43
Large Amount of News Items on a Page	42	62	31
Low Amount of News Items on a Page	29	27	79
Salaciousness of Content	17	90	28
Sordidness of Content	11	92	32
Obviously Syndicated Content	37	40	58

9.20.6 *Impact of Website Functionality, Third Party Functionality, Personalisation, and Advertising on the Perception of Bias*

Table 9-30 The impact of website functionality on the perception of bias.

Question: Have you ever detected bias in any of the following functionality of a news website?			
Input: Radio Button			
Answer	Yes	No	I Don't Know
Trending Articles or Content	88	15	32
Interaction facilities – Comment sections	87	12	36
‘Also read’, or recommended articles	63	41	31
Interaction facilities – voting	56	41	38
Breadcrumb trail or section heading	42	46	47
Main Navigation	41	58	36
Structure and Layout	25	71	39
Secondary Navigation	24	56	55
Website Functionality	12	88	35
Interaction facilities: register, signup, login	9	108	18

Table 9-31 The impact of third-party functionality on the perception of bias

Question: Have you ever detected bias in any social media facilities integrated with a news website?			
Input: Radio Button			
Answer	Yes	No	I Don't Know
Social media interaction facilities (like or share)	10	108	17
Social media (call to action, register, signup, follow)	11	112	12
Social media (sharing, views, live stream, and trending)	41	72	22
Promoted internal or external content	90	24	21

Table 9-32 The impact of personalisation on the perception of bias.

Question: Have you ever detected bias in any of the following personalised features of a news website?			
Input: Radio Button			
Answer	Yes	No	I Don't Know
Personalised content recommendation	71	61	3
Trending Social Media content	60	51	24
Personalised homepage	56	54	25
Visible personalised profiles	41	82	12

Table 9-33 The impact of advertising on the perception of bias.

Question: Have you ever detected bias in any advertising on a news website?			
Input: Radio Button			
Answer	Yes	No	I Don't Know
Advertising (Static or Animated)	113	21	1
Advertising (Video advertisement)	104	16	15
Advertising (popup or otherwise moving)	83	33	19
Advertising (Article or webpage background)	41	47	47
Advertising (Reputation Lending)	56	49	30
Advertising (Self-promotional)	69	32	34

9.21 Appendix U - Chapter Five: Experiment Two - Steps to Increase Validity

Like experiment one, six steps were taken to increase the validity of experiment two:

- Participants had to undertake two Instruction Tasks which included detailed directions on how to use the experiment interface including the VAS, see Figure 5-4. This helped to alleviate any issues with accidentally recording the wrong bias rating due to learning how to use the interface in the early stages of the experiment.
- Two attention questions measured participants' continued assiduousness to the task. Both attention questions used the same experiment interface with two news articles and websites not used in the experiment. This was to reduce the likelihood that lazy or inattentive participants would be alerted to the attention question. The first attention question was an article titled "Public Reminded of Dog Breed Rules" and included two pictures of dogs. The participant instruction at the bottom of the page contained the question "Is this News article about Dogs or Cats?" and "Please move the slider to the corresponding picture of the Dog or Cat", with the VAS adapted accordingly.
- After rating three distorted webpage/article combinations, participants were presented with all three side by side and given the opportunity to comparatively re-evaluate their submission by updating their perceived bias rating. This re-evaluation methodology had previously proven successful in the previous experiment (Spillane et al., 2017a). The results reported in this chapter are those re-evaluated ratings.
- To reduce the impact of confounding variables, un-emotive content was chosen, e.g. no articles about religion, sports teams, or the US election.
- Participants were recruited from the US to reduce the likelihood of them having preconceived strong emotions about any of the websites used in the experiment in case they recognised them.
- Participants' completion time was monitored to ensure that submissions from automated bots (<60 seconds) were discounted. Submissions which took less than 10 minutes were also discounted as pre-tests had shown it was impossible to complete the task properly in less than 15 minutes.

Further information on CAFE recommendations for increasing validity is available in section 3.6.5.3 of this thesis.

9.22 Appendix V – Chapter Five: Experiment Two – Reflective Questions

9.22.1 The Impact of Website Design and Webpage Aesthetics on the Perception of Bias and Professionalism

Table 9-34 The impact of design and aesthetics on perception of bias in the news article.

Question: Does a website’s design or a webpage’s aesthetic impact your perception of bias in the content of its articles?		
Input: Radio Button		
Answer	Responses	% of 405
Yes	128	31.60
No	99	24.44
Maybe	158	39.01
I Don’t Know	20	4.94

Table 9-35 The impact of design and aesthetics on perception of bias in the news organisation.

Question: Does a website’s design or a webpage’s aesthetic impact your perception of bias in the organisation behind the website?		
Input: Radio Button		
Answer	Responses	% of 405
Yes	154	38.02
No	104	25.68
Maybe	127	31.36
I Don’t Know	20	4.94

Table 9-36 The impact of supporting content on perception of bias in the news organisation.

Question: Do the links to other articles or other content in the same website influence your perception of bias in the News organisation?		
Input: Radio Button		
Answer	Responses	% of 405
Yes	210	51.85
No	85	20.99
Maybe	97	23.95
I Don’t Know	13	3.21

Table 9-37 The impact of design and aesthetics on perceived professionalism of news organisation.

Question: Do the links to other articles or other content in the same website influence your perception of the professionalism in the News organisation?		
Input: Quad Option Radio Button		
Answer	Responses	% of 405
Yes	267	65.92
No	59	14.56
Maybe	71	17.53
I Don’t Know	8	1.98

9.22.2 *The Features and/or Characteristics of a News Webpage that Impact Perceived Professionalism and Perceived Bias*

Table 9-38 The features and/or characteristics of a website’s design or a webpage’s aesthetics, or the article therein, that impact the perception of professionalism.

Question: Which features or characteristics of a news webpage or of the article therein, impact your perception of professionalism?			
Input: Multiple Choice			
Answer	Focus	Responses	% of 405
Low quality advertisements	Presentation	287	70.86
Emotive language	Text	280	69.14
Advertisements for cheap or low-quality products or services	Presentation	213	52.59
Lack of demonstrated research	Text	247	60.99
Too many advertisements	Presentation	263	64.94
Bad quality writing	Text	308	76.05
Advertisements which are too prominent or gaudy	Presentation	223	55.06
Unsupported claims in the article	Text	262	64.69
Low quality images	Presentation	65	16.05
Lack of balance in the article	Text	189	46.67
No information about the author	Presentation	126	31.11
Lack of objectivity in the article'	Text	220	54.32
No article information e.g. date stamp	Presentation	181	44.69
Lack of obvious history or research in the article	Text	215	53.09
Bad or broken alignment in the design	Presentation	182	44.94
Bad graphs or figures in the article	Text	176	43.46
Constant or oversized appeals to share content or connect via social media	Presentation	161	39.75
Obvious slant to support one individual position or argument	Text	220	54.32
Too much sponsored content	Presentation	215	53.09
Omitted facts from the article	Text	221	54.57
Discussion or comment facilities with signs of extreme views or lack of respect	Presentation	100	24.69
Bombastic headlines	Text	211	52.10
Bad colour schemes used in the design	Presentation	119	29.38
Lack of direct quotations of supporting data	Text	141	34.81
Advertisements which move, make sound or otherwise intrude on the user experience	Presentation	201	49.63
Use of words which have been deliberately chosen to give positive or negative impressions	Text	171	42.22
No links to other articles on the same topic or an obvious experience writing on the subject matter	Presentation	81	20.00
No links to other supporting articles	Text	92	22.72

Table 9-39 The features and/or characteristics of a website’s design or a webpage’s aesthetics, or the article therein, that impact the perception of bias.

Question: Which features or characteristics of a news webpage or of the article therein, impact your perception of bias?			
Input: Multiple Choice			
Answer	Focus	Responses	% of 405
Unsupported claims in the article	Text	244	60.25
Lack of objectivity in the article'	Text	228	56.30
Omitted facts from the article	Text	219	54.07
Obvious slant to support one individual position or argument	Text	215	53.09
Use of words which have been deliberately chosen to give positive or negative impressions	Text	213	52.59
Lack of balance in the article	Text	209	51.60
Bad or broken alignment in the design	Presentation	203	50.12
Low quality images	Presentation	199	49.14
Lack of obvious history or research in the article	Text	194	47.90
Emotive language	Text	192	47.41
Lack of demonstrated research	Text	191	47.16
Bombastic headlines	Text	190	46.91
Too much sponsored content	Presentation	186	45.93
Bad quality writing	Text	184	45.43
Bad colour schemes used in the design	Presentation	182	44.94
Too many advertisements	Presentation	175	43.21
Advertisements which are too prominent or gaudy	Presentation	174	42.96
Low quality advertisements	Presentation	167	41.23
Advertisements which move, make sound or otherwise intrude on the user experience	Presentation	162	40.00
Advertisements for cheap or low quality products or services	Presentation	159	39.26
Lack of direct quotations of supporting data	Text	158	39.01
Constant or oversized appeals to share content or connect via social media	Presentation	158	39.01
No information about the author	Presentation	131	32.35
Discussion or comment facilities with signs of extreme views or lack of respect	Presentation	120	29.63
No article information e.g. date stamp	Presentation	106	26.17
Bad graphs or figures in the article	Text	97	23.95
No links to other supporting articles	Text	96	23.70
No links to other articles on the same topic or an obvious experience writing on the subject matter	Presentation	84	20.74

9.22.3 Most Credible and Least Biased News Mediums

As part of the reflective questions in this experiment, participants were also asked which news mediums they found to be the most credible, see Table 9-40. While 29.63% of the participants chose television, 45.68% chose social media as the most credible. Unlike experiment one, social media was included as a separate option to the WWW as increasing numbers report getting their news from it in recent years. Since the 1960s, Roper reports have shown that television is the most credible news medium in the US (Gantz, 1981). However, Johnson and Kaye point out: “that people judge their preferred medium as the most credible” (Johnson & Kaye, 2000). Given that television has surpassed newspapers and radio as the most popular medium, it stands to reason that most studies, including Roper reports, find it to be the most credible medium. It also stands to reason that in an online crowdsourced experiment a large proportion of users would rate social media as being a credible medium. Despite this, it was surprising that newspapers were considered the least credible medium.

Table 9-40 Most credible new medium.

Question: Which news medium do you believe is the most credible?		
Input: Radio Button		
Answer	Responses	% of 405
Social Media	185	45.68
Television	120	29.63
WWW	59	14.57
Radio	21	5.19
Newspaper	20	4.94

Like experiment one, the reflective questions also asked participants which medium they considered to be the least biased. To the best of the authors knowledge, no similar questions have been asked in other studies. While this may seem surprising, searching “least biased medium” or similar in Google Scholar returns zero results (16/09/2018). It appears that most empirical studies or user surveys on bias tend to focus within a single medium, or on bias in multiple sources reporting of a single event. Thus, little or no research exists empirically comparing various mediums or asking users their impressions. Table 9-40 shows the results of the questions asking participants which medium they considered the least biased.

Like the previous question, social media was also included as a separate option to the WWW. The most interesting finding is the fact that despite newspapers being considered by over 40% of the participants to be the least biased, less than 4% considered social media to be the least biased. While it must be remembered that participants were crowdsourced online which may slant the results, this jars slightly with the results of the previous question. Newspapers, the least biased medium are also seen as the least credible.

Table 9-41 Least biased news medium.

Question: Which news medium do you believe is the least biased?		
Input: Radio Button		
Answer	Responses	% of 405
Newspaper	166	40.99
WWW	140	34.57
Television	55	13.58
Radio	28	6.91
Social Media	16	3.95

9.22.4 Dimensions of Credibility

Table 9-42 Dimensions of credibility most relied on when judging news on the WWW.

Question: When Judging the credibility of news on the WWW, which of the below dimensions of credibility do you rely on most? Please select all that apply. Input: Multiple Choice		
Answer	Responses	% of 405
Expertise	208	51.36
Trust	185	45.68
Depth of Coverage	233	57.53
Believability	150	37.04
Bias	217	53.58
Fairness	199	49.14
Honesty	234	57.78
Authoritative	67	16.54
Professional	241	59.51
Attractiveness	31	7.65
Interactive	21	5.19
Objective	219	54.07
Completeness	155	38.27
Accuracy	275	67.90
Integrity	203	50.12
Reputation	203	50.12
Successful	45	11.11
Reliability	195	48.15
Informative	204	50.37
Involving	28	6.91
Interesting	49	12.10
Factualness	223	55.06
Comprehension	103	25.43
Currency	20	4.94
Balance	108	26.67

D0
D1 Top

Progress on malaria deaths at serious risk without big boost in funding, UN warns

Real progress in driving down infections and deaths from malaria will be at serious risk if substantially more money is not paid by governments to the UN, according to the latest annual UN report on the epidemic.

Last year, more than 400,000 people, mostly small children and pregnant women, died from malaria. The UN spends the money it receives from governments around the world on measures to help combat the disease. This includes insecticide-impregnated bed nets to sleep under and effective drug treatments have brought the number of deaths down by nearly 30% worldwide in the five years from 2010 to 2015; the number of new cases over the same period is down by more than 20%.

But the report, published by the World Health Organisation, shows that some countries receive more funding than others and the UN now wants international governments to cough up more cash.

Some countries such as Sri Lanka and Kyrgyzstan have eliminated malaria and the WHO says the target of elimination in ten countries by 2020 will be met. Last year at least ten countries had 150 cases or fewer and nine more had between 150 and 1,000.

But the task of reducing the toll of malaria in the heavily endemic countries of sub-Saharan Africa, which have 90% of cases and 92% of deaths, is hard and needs more resources, says the report.


"We are definitely seeing progress," said Dr Pedro Alonso, director of the WHO Global Malaria Programme. "But the world is still struggling to achieve the high levels of programme coverage that are needed to beat this disease."

NewsCom Sign In News Sport Weather Shop Earth Travel More Search

NEWS


[Home](#) [Video](#) [World](#) [UK](#) [Business](#) [Tech](#) [Science](#) [Magazine](#) [Entertainment & Arts](#) [Health](#) [World News TV](#) [More](#)

[Asia](#) [China](#) [India](#)



Progress on malaria deaths at serious risk without big boost in funding, UN warns

01 June 2017 | Asia f t e p s



Malaria killed more than 400,000 people worldwide in 2016 JOSEPH MOTEN

Real progress in driving down infections and deaths from malaria will be at serious risk if substantially more money is not paid by governments to the UN, according to the latest annual UN report on the epidemic.

Last year, more than 400,000 people, mostly small children and pregnant women, died from malaria. The UN spends the money it receives from governments around the world on measures to help combat the disease. This includes insecticide-impregnated bed nets to sleep under and effective drug treatments have brought the number of deaths down by nearly 30% worldwide in the five years from 2010 to 2015; the number of new cases over the same period is down by more than 20%.

But the report, published by the World Health Organisation, shows that some countries receive more funding than others and the UN now wants international governments to cough up more cash.

Some countries such as Sri Lanka and Kyrgyzstan have eliminated malaria and the WHO says the target of elimination in ten countries by 2020 will be met. Last year at least ten countries had 150 cases or fewer and nine more had between 150 and 1,000.

But the task of reducing the toll of malaria in the heavily endemic countries of sub-Saharan Africa, which have 90% of cases and 92% of deaths, is hard and needs more resources, says the report.

"We are definitely seeing progress," said Dr Pedro Alonso, director of the WHO Global Malaria Programme. "But the world is still struggling to achieve the high levels of programme coverage that are needed to beat this disease."

Related Topics

[Asia](#) [Pakistan](#) [India](#)

Share this story About sharing

[Email](#) [Facebook](#) [Twitter](#) [LinkedIn](#) [Pinterest](#)

More on this story

First Real-World Test for Malaria Vaccine
15 March 2015

Medical research centre frozen in time
8 November 2014

Malaria: Kenya, Ghana and Malawi get first vaccine
20 December 2016

Asia


UN urges Thailand to amend lese-majeste law
19 June 2017 | Asia

China cracks down on online moneylenders
19 June 2017 | China


Dried lizard penis sold as tantric root
19 June 2017 | India

More Videos from NewsCom


Recommended by Outlook




The Indian video challenging 'sury bride' stereotypes




London fire: How one Grenfell Tower resident escaped




Georgia prisoners Rowe and Dutoose caught after massive manhunt




Outrage after rapist jailed with victim




Spalding murders: Daughter Kim Edwards named as killer



Is refusing terrorists funerals a deterrent?



Chicago goes high-tech in search of answers to gun crime surge




The benign menace of Bollywood's cult 'monster'


Figure 9-10 The top half of the D0 and D1 webpage/article combinations used in experiment three.

D1 Bottom


Elsewhere on NewsCom
Recommended by Outlook




NewsCom News
Newspaper headlines: May's Brexit plan and S&S pose as homeless




NewsCom Travel
The sheikh who saved my family




NewsCom News
Call for sex book ban at temples lined with erotic art



NewsCom Culture
The man who ignored a tornado



NewsCom News
Qatar crisis: Have Saudis gone too far?




NewsCom Future
Why printers add secret tracking dots

From intersex stigma to gospel star


Most Read

- 1 **Finebury Park attack suspect named as Darren Osborne**
- 2 **Brian Cant, Play School presenter, dies at 83**
- 3 **Croatia football captain Modric faces injury inquiry**
- 4 **Champs-Élysées: Car 'deliberately hit' police van in Paris**
- 5 **Amanda Holden's 'topless dress' escapes Ofcom investigation**
- 6 **Carrie Fisher 'had cocaine' in her system when she died**
- 7 **Brexit negotiations: David Davis says 'promising start' made**
- 8 **Grenfell Tower fire: Police identify five victims of blaze**
- 9 **Syria conflict: Russia issues warning after US coalition downs jet**
- 10 **M&S says labelling avocados with lasers is more sustainable**


ADVERTISEMENT




ADVERTISEMENT




From Around the Web
Promoted content by Outlook




If you own a computer you must try this game!
Thomas Kingston at W...




Top 11 Richest Cricketers In The World
RichieBartlett@r...




Free Pattern Friday: Creativity Takes Flight
Cathy | Leah H. Ma...



12 Knitting Patterns That Pair Perfectly With Hand-Dyed...
Cathy | Leah H. Ma...



Small Space Design Inspiration from Across the Pond
Lonny Magazine



This game will keep you up all night!
Manga

Ads by Google

British Expat in America?
Avoid Losing a Large Part of your UK pension. Move your pension. Get your Free Review! Forth Capital

Online LLM in 1 Year
Advance your career with an Online Master of Laws. Learn more. Univ. of Cumbria at RUC

Online LLM in 1 Year
Advance your career with an Online LLM in Int. Commercial Law. Request more information Univ. of Salford @ RUC

Jobs in America
11 million people trust us with their job search - upload your CV & join today. cv-library.ie

NewsCom Services

On your mobile
 On your connected tv
 Get news alerts
 Contact NewsCom News

Explore NewsCom

Home	News	Sport	Weather	Shop	Earth
Travel	Capital	Culture	Autos	Future	TV
Radio	Children's NC	NCables	Food	Wonder	Stress
Music	Arts	Make It Digital	Tester	Nature	Local

[Terms of Use](#)
 [About NewsCom](#)
 [Privacy Policy](#)
 [Cookies](#)
 [Accessibility Help](#)
 [Parental Guidance](#)
 [Contact NewsCom](#)
 [Advertise with us](#)
 [Ad choices](#)


Copyright © 2017 NewsCom. The NewsCom is not responsible for the content of external sites. Read about our approach to external linking.

Figure 9-11 The bottom half of the D0 and D1 webpage/article combinations used in experiment three.

D2 Top

D3 Top

Advertisement



ROLLS-ROYCE
WRAITH

NewsCom Sign In News Sport Weather Shop Earth Travel More Search


NEWS

Home Video World UK Business Tech Science Magazine Entertainment & Arts Health World News TV More

Asia China India

Progress on malaria deaths at serious risk without big boost in funding, UN warns

01 June 2017 Asia



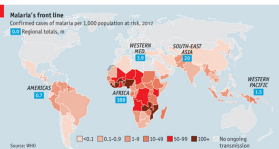
Malaria killed more than 400,000 people worldwide in 2016

Real progress in driving down infections and deaths from malaria will be at serious risk if substantially more money is not paid by governments to the UN, according to the latest annual UN report on the epidemic.

Last year, more than 400,000 people, mostly small children and pregnant women, died from malaria. The UN spends the money it receives from governments around the world on measures to help combat the disease. This includes insecticide-impregnated bed nets to sleep under and effective drug treatments have brought the number of deaths down by nearly 20% worldwide in the five years from 2010 to 2015; the number of new cases over the same period is down by more than 20%.

But the report, published by the World Health Organisation, shows that some countries receive more funding than others and the UN now wants international governments to cough up more cash.

Some countries such as Sri Lanka and Kyrgyzstan have eliminated malaria and the WHO says the target of elimination in ten countries by 2020 will be met. Last year at least ten countries had 150 cases or fewer and nine more had between 150 and 1,000.



Malaria's front line
Confirmed cases of malaria per 1,000 population at risk, 2015

But the task of reducing the toll of malaria in the heavily endemic countries of sub-Saharan Africa, which have 90% of cases and 92% of deaths, is hard and needs more resources, says the report.

"We are definitely seeing progress," said Dr Pedro Alonso, director of the WHO Global Malaria Programme. "But the world is still struggling to achieve the high levels of programme coverage that are needed to beat this disease."

Related Topics: Asia Pakistan India

Share this story about sharing

James E. Jones is the Chief Environmental Correspondent for NewsCom. A graduate of Princeton he previously held the position of senior international development officer with the UN.

Correction: The 10th paragraph of this article was amended at 14:54 on June 2nd 2017 to show data was deaths were down by 30% rather than by 33% as previously stated

More on this story

First Real World Test for Malaria Vaccine
15 March 2015

Medical research centre frozen in time
8 November 2014

Malaria: Kenya, Ghana and Malawi get first vaccine
20 December 2016

Asia

UN urges Thailand to amend lese-majeste law
19 June 2017 Asia

China cracks down on online moneylenders
19 June 2017 China

Dried lizard penis sold as tantric root
19 June 2017 India

More Videos from NewsCom

Top Stories

London mosque attack suspect named
38 minutes ago

Brexit talks begin
4 hours ago

Russian warning after US downs Syrian jet
27 minutes ago

Most Read

Finbury Park attack suspect named as Darren Osborne
2

Brian Cant: Play School presenter, dies at 83

Croatia football captain Modric faces perjury inquiry

Champs-Elysees: Car 'sabermetally hits' police van in Paris

Amanda Holden's 'topless dress' escapes Ofcom investigation

Carris Fisher 'had cocaine' in her system when she died

Brexit negotiations: David Davis says 'promising start' made

Grenfell Tower fire: Police identify five victims of blaze

Syria conflict: Russia issues warning after US coalition downs jet

M&S says labelling avocados with isears is more sustainable

Features

Coffee under threat

Macron's conquest is complete - what now?

UN urges Thailand to amend lese-majeste law

The hijabi ballerina

Allowing mums in jail to hug their children

The Tunisian women who want to be virgins again

Chicago goes high-tech in search of answers to gun crime surge

The benign menace of Bollywood's cult 'monster'

Global News Network

NewsCom News Sport Weather Shop Earth Travel More Search

NEWS


Home Video World UK Business Tech Science Magazine Entertainment & Arts Health World News TV More

UP TO \$200 BONUS FOR NEW CUSTOMERS bet365 JOIN NOW

Terms & Conditions Apply Gamble Responsibly

Progress on malaria deaths at serious risk without big boost in funding, UN warns

01 June 2017 Asia



Malaria killed more than 400,000 people worldwide in 2016

Real progress in driving down infections and deaths from malaria will be at serious risk if substantially more money is not paid by governments to the UN, according to the latest annual UN report on the epidemic.

Last year, more than 400,000 people, mostly small children and pregnant women, died from malaria. The UN spends the money it receives from governments around the world on measures to help combat the disease. This includes insecticide-impregnated bed nets to sleep under and effective drug treatments have brought the number of deaths down by nearly 20% worldwide in the five years from 2010 to 2015; the number of new cases over the same period is down by more than 20%.

But the report, published by the World Health Organisation, shows that some countries receive more funding than others and the UN now wants international governments to cough up more cash.

Some countries such as Sri Lanka and Kyrgyzstan have eliminated malaria and the WHO says the target of elimination in ten countries by 2020 will be met. Last year at least ten countries had 150 cases or fewer and nine more had between 150 and 1,000.

But the task of reducing the toll of malaria in the heavily endemic countries of sub-Saharan Africa, which have 90% of cases and 92% of deaths, is hard and needs more resources, says the report.

"We are definitely seeing progress," said Dr Pedro Alonso, director of the WHO Global Malaria Programme. "But the world is still struggling to achieve the high levels of programme coverage that are needed to beat this disease."

Share with your friends!

Help us spread the word by sharing NewsCom with your friends and followers on social media!

Share on Facebook Share on Twitter G+ Pinterest Email

Real progress in driving down infections and deaths from malaria will be at serious risk if substantially more money is not paid by governments to the UN, according to the latest annual UN report on the epidemic.

Last year, more than 400,000 people, mostly small children and pregnant women, died from malaria. The UN spends the money it receives from governments around the world on measures to help combat the disease. This includes insecticide-impregnated bed nets to sleep under and effective drug treatments have brought the number of deaths down by nearly 20% worldwide in the five years from 2010 to 2015; the number of new cases over the same period is down by more than 20%.

But the report, published by the World Health Organisation, shows that some countries receive more funding than others and the UN now wants international governments to cough up more cash.

Features

Coffee under threat

Macron's conquest is complete - what now?

UN urges Thailand to amend lese-majeste law

The hijabi ballerina

Allowing mums in jail to hug their children

The Tunisian women who want to be virgins again

Chicago goes high-tech in search of answers to gun crime surge

The benign menace of Bollywood's cult 'monster'

More Videos from NewsCom

Recommended by Outbrain

The Indian video challenging 'shy bride' stereotypes

London fire: How one Grenfell Tower resident escaped

Georgia prisoners Rovo and Dubose caught after massive manhunt

Outrage after rapist jailed with victim

Spalding murderer: Daughter Kim Edwards named as killer

Is re-living terrorise funerals a deterrent?


The benign menace of Bollywood's cult 'monster'

Figure 9-12 The top half of the D2 and D3 webpage/article combinations used in experiment three.


D2 Bottom

D3 Bottom


Advertisement by Outlook



The Indian video challenging 'why bride' stereotypes




London fire: How one Grenfell Tower resident escaped




Georgian prisoners Rowe and Dubose caught after massive manhunt

Elsewhere on NewsCom


Recommended by Outlook




NewsCom News
Newspaper headlines: May's Brexit plan and S&B pose as homeless



NewsCom Travel
The sheikh who saved my family



NewsCom News
Call for sex book ban at temples lined with erotic art



NewsCom Culture
The man who ignored a tornado

NewsCom Services


On your mobile On your connected tv Get news alerts Contact NewsCom News

Explore NewsCom

Home	News	Sport	Weather	Shop	Earth
Travel	Capital	Culture	Autos	Future	TV
Radio	Children's NC	NCubates	Food	Wonder	Blissies
Music	Arts	Make It Digital	Tester	Nature	Local

Terms of Use About NewsCom Privacy Policy Cookies Accessibility Help Parental Guidance Contact NewsCom Advertise with us All rights reserved.
Copyright © 2017 NewsCom. The NewsCom is not responsible for the content of external sites. Read about our approach to external linking.

Sponsored Financial Content




MarketViews


Top 10 Trading Mistakes with CFDs: free 5 page guide - Capital at Risk

Why Invest in Currency Trader's First Pick


Spponsored stories



Why Doctors No Longer Prescribe Adderall: Dr. N. D. D.



I Think You Should Use Every Inch of Your PC...




The Game You Can't Create as If From Land

YOUNGER FACE & EYES IN MINUTES


DERMATACE

Elsewhere on NewsCom


Recommended by Outlook




NewsCom News
Newspaper headlines: May's Brexit plan and S&B pose as homeless




NewsCom Travel
The sheikh who saved my family



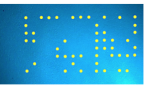
NewsCom News
Call for sex book ban at temples lined with erotic art



NewsCom Culture
The man who ignored a tornado



NewsCom News
Qatar crisis: Have Saudis gone too far?



NewsCom Future
Why printers add secret tracking dots

From Around the Web

If you own a computer you must try this game!

Top 11 Richest Cricketers in the World

Free Pattern Friday: Creativity Takes Flight

12 Knitting Patterns That Pair Perfectly With Hand-Dyed

Small Space Design Inspiration from Across the Pond

This game will keep you up all night!


Ads by Google

British Expert in America? Avoid Losing a Large Part of your UK Pension. Move your pension. Get your Free Review Form: Capital Online LLM in 1 Year Advance your career with an Online Master of Laws. Learn more. Univ. of Cumbria at BNC


Online LLM in 1 Year Advance your career with an Online LLM in Int. Commercial Law. Request more information Univ. of Eastern @ BNC

Jobs in America 11 million people trust us with their job search - upload your CV & job today. on-library


promoted links from around the web Recommended by Outlook




World's Most Expensive Home Hits Market for £1.1m




Turkey: In the Crossfire - ANTE Reportage




Teen Vanishes in Florida, But 25 Years Later, Her Family...




Do This Before Bet To Regain You Have All Right...




Donald Trump's St. Martin Vacation Compound Gets \$11 Million Price Cut



Boost your menu boards with these 10 pro tips!



The most addictive game of the year! Play with 14 million players now!



The Game Lets You Create an Elf Fantasy Land!

NewsCom Services

On your mobile On your connected tv Get news alerts Contact NewsCom News

Explore NewsCom

Home	News	Sport	Weather	Shop	Earth
Travel	Capital	Culture	Autos	Future	TV
Radio	Children's NC	NCubates	Food	Wonder	Blissies
Music	Arts	Make It Digital	Tester	Nature	Local

Figure 9-13 The bottom half of the D2 and D3 webpage/article combinations used in experiment three.

9.24 Appendix X – Chapter Five and Six: Experiment Two and Three - Participant Sample Profile

405 submission, 51.4% Male, 47.2% Female, 0.7% Other, and 0.7% Prefer not to say. The Mean age was 36.48 years old. 34% of participants were <30.

Participants were spread across the US with 44 states and the District of Columbia represented. California (49), Texas (25) and New York (20) had the most representation.

Education level was high with 57% having completed a four-year bachelor degree or higher.

Occupation had 4.9% reporting home parent, 13.8% student, 9.6% unemployed and 9.1% other. The other participants were heavily weighted towards professional services with 55.5% reporting as professional, semi-professional or skilled professional. 15% report their occupation as skilled manual, semi-skilled manual or unskilled manual.

Income level, 38.8% report earning less than \$29,999 per annum, 29.4% between \$30,000 and \$59,999, 17.5% between \$60,000 and \$89,999, 8.2% between \$90,000 and \$119,999. The remainder, 6.1% report incomes greater than \$120,000.

Political Leaning, 56% of participants reported being Strong, Not So Strong or Independent Leaning Liberal. 24.7% report being Strong, Not So Strong, or Independent Leaning Conservative. 15.1% report Independent while 2.2% report Other and 2% report Don't Know.

Table 9-43 Sample profile breakdown of the participants in experiments two and three.

N = 405	Category	Frequency	Percentage
States	44 US States and District of Columbia	405	100%
Gender	Male	208	51.4%
	Female	191	47.2%
	Other	3	0.7%
	Prefer not to say	3	0.7%
Age (Mean 36.48)	18 - 30	151	37.28%
	31 - 40	132	32.59%
	41 - 50	53	13.09%
	51 - 60	45	11.11%
	61 - 70	19	4.69%
	71 - 80	4	0.98%
	81 - +	1	0.25%
Education	Elementary school only	1	0.25%
	Some high school, but did not finish	5	1.23%
	Completed high school	50	12.35%
	Some college, but did not finish	77	19.01%
	Two-year college degree / A.A / A.S	41	10.12%
	Four-year college degree / B.A. / B.S	134	33.09%
	Some graduate study	26	6.42%
	Completed Masters or professional degree	47	11.60%
	Advanced Graduate study or Ph.D.	22	5.43%
	Other	2	0.49%
Internet Usage	Internet less than one hour a day	1	0.25%
	Internet 1 – 2 hours a day	11	2.72%
	Internet 2 - 4 hours a day	72	17.78%
	Internet 4 - 6 hours a day	111	27.41%
	Internet 6 - 8 hours a day	84	20.74%
	Internet 8 plus hours a day	126	31.11%
Occupation	Stay at home Parent	20	4.94%
	Student	56	13.82%
	Unemployed	39	9.63%
	Professional	111	27.41%
	Semi-professional	52	12.84%
	Skilled professional	62	15.31%
	Skilled manual	17	4.20%
	Semi-skilled manual	7	1.73%
	Unskilled manual	4	0.99%
	Other	37	9.13%
Political Leaning	Strong Liberal	101	24.94%
	Not so strong Liberal	56	13.83%
	Independent leaning Liberal	70	17.28%
	Independent	61	15.06%
	Independent leaning Conservative	27	6.67%
	Not so strong Conservative	45	11.11%
	Strong Conservative	28	6.91%
	Other	9	2.22%
	Don't know	8	1.97%