

5-1-2020

Three essays on malicious consumer deviance: The creation, dissemination, and elimination of misleading information

Tyler Hancock

Follow this and additional works at: <https://scholarsjunction.msstate.edu/td>

Recommended Citation

Hancock, Tyler, "Three essays on malicious consumer deviance: The creation, dissemination, and elimination of misleading information" (2020). *Theses and Dissertations*. 4726.
<https://scholarsjunction.msstate.edu/td/4726>

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

Three essays on malicious consumer deviance: The creation, dissemination, and elimination of
misleading information

By

Tyler Hancock

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Business Administration (Marketing)
in the Department of Marketing, Quantitative Analysis, and Business Law

Mississippi State, Mississippi

May 2020

Copyright by
Tyler Hancock
2020

Three essays on malicious consumer deviance: The creation, dissemination, and elimination of
misleading information

By

Tyler Hancock

Approved:

Frank G. Adams
(Major Professor)

Michael J. Breazeale
(Major Professor)

Jason E. Lueg
(Committee Member)

Kevin J. Shanahan
(Committee Member)

Laura Allen
(Committee Member)

Nicole Ponder
(Graduate Coordinator)

Sharon L. Oswald
Dean
College of Business

Name: Tyler Hancock

Date of Degree: May 1, 2020

Institution: Mississippi State University

Major Field: Business Administration (Marketing)

Major Professors: Frank G. Adams and Michael J. Breazeale

Title of Study: Three essays on malicious consumer deviance: The creation, dissemination, and elimination of misleading information

Pages in Study: 273

Candidate for Degree of Doctor of Philosophy

With the explosion of social media, consumers are gaining control in social reach and can utilize online platforms to create and share misleading information when doing so helps to meet an end.

This dissertation, consisting of three separate essays, represents an attempt to address how misleading information is created, how it is disseminated, and how it can be eliminated.

Essay One (Chapter 2) uses a mixed-method approach to explore the Dark Triad, proactivity, and vigilantism in driving self-created misleading information sharing. Additionally, this essay introduces a dual-process model of inoculation theory to the marketing and consumer literature that shows how consumers autoinoculate when building justification to engage in malicious behavior. This process includes both automatic and analytical components that initiate a Negative Cascade.

Without a larger number of posts, these initial messages may be overlooked. However, herd inoculation can develop when a message begins to sway larger groups. Essay Two (Chapter 3) determines that authentic messages from the original poster are most believable and most likely to initiate a Negative Cascade. This confirmation through mere exposure can then initiate herd inoculation as it flows to other consumers and develops further credibility. The implicit

bystander effect is active when in the presence of larger groups. Findings suggest herd inoculation may go unbroken since posters exposed to a positive counter-cascade are less likely to both participate in a forum and post positive messages.

Essay Three (Chapter 4) shows that when a consumer shares a message that develops into a Negative Cascade, additional effort is required to halt the consumer herd inoculation. The studies uncover the need for an overt response from the original poster to stop future sharing of misleading information and the role of brand-enacted quarantines in the prevention of the autoinoculation of consumer vigilantes.

This dissertation shows how one message can become a much bigger problem for a brand when misinformation spreads. Insights within the dissertation provide numerous outlets for future research and numerous tools and recommendations for both academics and practitioners that hope to understand how misleading information is created, disseminated, and can be eliminated.

Keywords: inoculation, vigilantism, Dark Triad, availability cascades, bystanders, misinformation-sharing

DEDICATION

I dedicate this dissertation to my wife, Yessica, my parents, Doug and Sandra, and all of my friends and family who have supported me throughout my pursuit of a career in academia. Without your support, I would not be the person that I am today.

Yessica, when I first met you, I talked about my interest in pursuing a Ph.D. trying to impress you. You always make me want to learn and be the best version of myself that I can be. Whether it be my time in sales, consulting, or teaching, you have helped me to grow. Every day you provide me with encouragement and strength and help me to be patient and persistent when faced with challenges. I cannot put into words how grateful I am for all of the support you have provided and the sacrifices you have made to help me achieve this accomplishment. You mean the world to me, “and the world is worth a lot these days.”

Mom and Dad, what can I say? I guess I “knew what I had to do.” Thank you so much for pushing me to work hard and aim high for the things that are important to me. Thank you for preventing me from taking a break from school, and thank you for always supporting me every step of the way. From your constant support to words of encouragement, I am genuinely grateful for everything that you do. I am truly blessed by all of the love and support you have given me.

ACKNOWLEDGEMENTS

I want to thank and show appreciation to all of the people who both supported and aided in the development of this dissertation. A scholar is only as good as his mentors, and I am fortunate to have met many during my time at Mississippi State University. I am truly grateful for my co-chairs – Drs. Frank Adams and Michael Breazeale. To Dr. Frank Adams, even since we met in my interview, you have shown yourself to be an outlet for honesty and integrity. You have continued to prove yourself as a benefactor for not only me but also all of the other doctoral students in the program. You have helped me to develop as an academic and have ultimately made me a better scholar. You have set an example that I hope to embody throughout my career. Thank you for always having a “second” to touch-base about so many random topics. To Dr. Michael Breazeale, thank you for helping me to chase ideas that are fun and exciting to me. Since meeting with you and discussing the earlier concepts for my dissertation, you provided feedback, support, and enthusiasm. You have also helped other doctoral students and me to build confidence as both researchers and teachers. I hope to bring this continued enthusiasm into future projects and the classroom for years to come. I can confidently say that I am a much better writer and researcher due to our work together.

I would also like to extend my appreciation to my dissertation committee. Dr. Jason Lueg, thank you for helping me to learn about the good, bad, and ugly of the research process. You gave me a chance to contribute and build projects, and I am genuinely grateful for the support you have provided. I appreciate your investment in me as a scholar, and I know that the

knowledge you have instilled will serve me well throughout my career. Dr. Kevin Shanahan, even on the most stressful days, I could count on you to make us laugh and smile. Since day one, you have supported my ideas, and you have instilled in me a curiosity for exciting and novel ideas. I hope to one-day work and collaborate with both academics and students in both conferences and institutional settings like you do. Dr. Laura Allen, I am sincerely grateful for the support you have provided and the interesting perspectives you have conveyed. Your class helped to drive the development of my dissertation, and I hope to implement these techniques in my future research.

The faculty at Mississippi State provided a collaborative and inclusive environment that made me feel like I am at home. To Dr. Melissa Moore and the entire MQABL family, I cannot imagine being in a more supportive and collegial department. I am in awe of the support that has been provided throughout my time in the program. Whether it be through discussing ideas, developing research topics, or sharing advice and support in the classroom, I have been thoroughly supported. I hope to continue this culture of collegiality at my future institutions. To my doctoral student colleagues Christian Barney, Brett Kazandjian, Haley Hardman, Kavi Nambisan, and Jutong Wen, thank you for the awesome lunch discussions and for making the doctoral process that much more enjoyable. To my doctoral program mentors, Jennifer Stevens and Stacie Waites, thank you for providing me the confidence to grow as both a scholar and coauthor. I want to thank Dr. Mike Wittmann, Dr. Jamye Foster, Dr. Leisa Flynn, Dr. Melinda McLelland, and all of the teachers, professors, and colleagues that have encouraged me to pursue this path and helped provide the support for me to succeed.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	x
LIST OF FIGURES	xii
CHAPTER	
I. INTRODUCTION	1
Essay One: Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information	2
Essay Two: The Herd and the Bystander: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments	3
Essay Three: Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information.....	4
Organization of the Dissertation.....	5
II. ESSAY ONE	7
Introduction	8
Literature Review	12
Qualitative Research.....	17
Study 1: Netnographic content analysis	17
Netnographic Content Analysis Research Design.....	18
LIWC	21
Findings Related to Online Information Sharing and Trolling Behavior.....	24
Motivations for Sharing Misleading Information.....	24
Traits that Drive Misleading Online Information Sharing	33
Machiavellianism	34
Narcissism	37
Psychopathy.....	38
Application of Themes to Forum Discussing Pricing Mistake	41
Proactivity in Pursuing Opportunities	42
Vigilantism	46
Justification.....	49
Inoculation Theory	56

Conceptual Development of the Model.....	60
The Dark Triad	60
Machiavellianism	61
Narcissism	62
Psychopathy.....	62
Proactivity in Pursuing Opportunities	63
Vigilantism	64
Sharing Self-Created Misleading Information	64
Dark Triad and Proactivity	65
Proactivity and Vigilantism.....	66
The Dark Triad, Proactivity, and Vigilantism	67
Justification (Autoinoculation).....	68
Quantitative Research.....	70
Study 2: Testing the Model	70
Measures Pretest.....	70
Model Pretest.....	73
Sample and Procedure	74
The Online Disinhibition Effect	76
Confirmatory Factor Analysis	76
Structural Equation Modeling	77
The Dark Triad Formative Model	78
The Dark Triad and Proactivity.....	79
Proactivity and Vigilantism.....	80
Mediation.....	80
Dark Triad=> Proactive => Vigilantism	80
Proactive=> Vigilantism => Sharing Self-created Misleading Information	82
Dark Triad => Proactive => Vigilantism=> Sharing Self-created Misleading Information	82
Moderated Mediation	83
Vigilantism and Justification.....	83
The Dark Triad, Vigilantism, and Justification	84
Discussion and Implications.....	85
Theoretical Implications.....	87
Managerial Implications.....	88
Limitations and Future Research.....	91
Conclusion.....	92
III. ESSAY TWO	93
Introduction	94
Theoretical Framework	98
Inoculation Theory: A Dual-Process Model.....	98
Herd Inoculation.....	99
Availability Cascades	101
Believability of the Message	103
Justification.....	104

Likelihood to Spread the Message	105
Believability and Justification	106
Negative and Positive Cascades	107
Study 1: Message Type	109
Message Type Believability	109
LIWC Treatments	110
MANOVA	111
Procedure	112
Results	116
Study 2	117
Pretest	117
Procedure	118
Mediation	124
Message Type => Believable => Justifiable=> Spreading Message.....	125
Negative Cascade/ No Cascade Collection	125
Positive Cascade/ No Cascade Collection	125
Moderated Mediation	126
Negative Cascade	126
Positive Cascade	127
Discussion.....	128
Consumer Bystanders	130
Believability of the Message	132
Justification.....	133
Posting a Positive Message	134
Posting a Negative Message.....	135
The Bystander Effect.....	135
Study 3: The Bystander Effect	137
Sample and Procedure	137
Confirmatory Factor Analysis	140
Invariance	142
Structural Equation Modeling	142
Results	143
Message Type and Believability	143
Believability and Justification in Forum Tone	143
Justification in Forum Tone and Posting Positive Information.....	144
Justification in Forum Tone and Posting Negative Information	145
Mediation.....	146
Message Type => Believable => Justifiable=> Post Positive	147
Negative Cascade	147
Positive Cascade.....	148
Mediation Comparison	149
Moderated Mediation	149
Discussion.....	151
General Discussion and Implications	152
Limitations and Future Research.....	155

Conclusion.....	156
IV. ESSAY THREE.....	157
Introduction	158
Background.....	164
Consumer Advocacy	164
Consumer Vigilantism.....	165
Consumer Deviance.....	166
Inoculation Theory	167
Herd Inoculation.....	168
Availability Cascades	169
Hypothesis Development.....	171
Believability of the Message	171
Justification in Sharing the Message	172
Believability and Justification	173
Halting Herd Inoculation.....	174
Vigilante Responses	176
Discreet Vigilante Responses.....	177
Overt Vigilante Responses	179
Study 1: Halting the Herd Inoculation.....	181
Method.....	181
Results	185
Mediation.....	185
Believable => Justifiable=> Spreading Message	186
Moderated Mediation	186
Believable =>Justification*Vigilante Response =>Share the Message	186
Discussion.....	188
Misinformed Vigilantes.....	190
Quarantining the Vigilante	191
Vigilantism Propensities.....	192
Justification for a Decision.....	193
Implementing a Quarantine	194
Study 2.....	196
LIWC Treatments	197
Brand Response	198
ANOVA.....	200
Procedure	201
Results	203
Vigilantism*Justification.....	203
Implementing the Quarantine	204
Vigilantism*Justification* Brand Response.....	204
Discussion.....	205
General Discussion and Implications	206
Theoretical Implications	208
Managerial Implications.....	210

Limitations and Future Research.....	212
Conclusion.....	214
V. OVERVIEW AND CONCLUSION	215
Essay One: Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information.....	216
Essay Two: The Herd and the Bystander: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments.....	217
Essay Three: Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information.....	219
Synthesis and Final Remarks.....	220
REFERENCES	223
APPENDIX	
A. ESSAY ONE	245
Memos	246
1. Present – 12/2018	246
2. Initial Exposure – 02/2016	246
3. Updates from Promotion and Distribution Seminar – 10/2016.....	247
4. Updates from Qualitative Seminar – 08/2017	247
5. Updates from Consumer Behavior Seminar – 10/2017.....	248
6. Second-year paper review – 02/2018	249
7. Research Design Seminar – 5/2018.....	249
8. Casting a broader net to numerous Forums – 8/2018.....	250
LIWC Output Overview	252
Scales	254
Scenarios.....	257
B. ESSAY TWO	258
Company Description for Collection.....	259
MANOVA Treatments	259
C. ESSAY THREE.....	265
Vigilante Response Treatments.....	266
Brand Quarantine Treatments.....	268
Essay Three: Study 2 Interaction Visualizations.....	270
D. CONSUMER INOCULATION PROCESS	272
Consumer Inoculation Process and Recommended Treatments	273

LIST OF TABLES

Table 2.1	Hypotheses	70
Table 2.2	CFA Results, t-values, and Composite Reliability	72
Table 2.3	Sample Description	75
Table 2.4	Means, Standard Deviations, AVEs, and Shared Variance	77
Table 2.5	Hypothesis Results	81
Table 3.1	Study 1 Sample Description	113
Table 3.2	CFA Results, t-values, and Cronbach's α	115
Table 3.3	Means, Standard Deviations, AVEs, and Shared Variance	115
Table 3.4	ANOVA Results	116
Table 3.5	Study 2 Sample Description	119
Table 3.6	CFA Results, t-values, and Cronbach's α	121
Table 3.7	Means, Standard Deviations, AVEs, and Shared Variance	122
Table 3.8	Studies 2a and 2b Results	124
Table 3.9	Study 3 Sample Description	138
Table 3.10	CFA Results, t-values, and composite reliability	141
Table 3.11	Means, Standard Deviations, AVEs, and Shared Variance	142
Table 3.12	Study 3 Results	146
Table 3.13	Mediation Analysis	146
Table 4.1	Study 1 Sample Description	182
Table 4.2	CFA Results, t-values, and Cronbach's α	184

Table 4.3	Means, Standard Deviations, AVEs, and Shared Variance.....	185
Table 4.4	Hypothesis Results	185
Table 4.5	Study 2 Sample Characteristics	199
Table 4.6	CFA Results, t-values, and Cronbach's α	202
Table 4.7	Means, Standard Deviations, AVEs, and Shared Variance:	203
Table 4.8	Study 2 Results	205

LIST OF FIGURES

Figure 2.1	LIWC Summary Outputs Example	22
Figure 2.1	Poster 1 Summary.....	26
Figure 2.2	Poster 2 Summary.....	27
Figure 2.3	Poster 3 Summary.....	28
Figure 2.4	Poster 4 Summary.....	30
Figure 2.5	Poster 5 Summary.....	31
Figure 2.6	Poster 6 Summary.....	35
Figure 2.7	Poster 7 Summary.....	35
Figure 2.8	Poster 8 Summary.....	36
Figure 2.9	Poster 9 Summary.....	38
Figure 2.10	Poster 10 Summary.....	39
Figure 2.11	Figure 2.10 (continued)	40
Figure 2.12	Poster 12 Summary.....	44
Figure 2.13	Poster 13 Summary.....	44
Figure 2.14	Poster 14 Summary.....	45
Figure 2.15	Poster 15 Summary.....	46
Figure 2.16	Poster 16 Summary.....	48
Figure 2.17	Poster 17 Summary.....	49
Figure 2.18	Poster 18 Summary.....	50
Figure 2.19	Poster 19 Summary.....	51

Figure 2.20 Poster 20 Summary.....	52
Figure 2.21 Poster 21 Summary.....	53
Figure 2.22 Poster 22 Summary.....	54
Figure 2.23 Poster 23 Summary.....	54
Figure 2.24 Conceptual Model	69
Figure 2.25 Full Model and Results.....	85
Figure 3.1 Conceptual Models: Studies 2a and 2b	109
Figure 3.2 MANOVA	112
Figure 3.3 Study 2a and 2b Model Results	128
Figure 3.4 Study 3 Conceptual Model	137
Figure 4.1 Study 1 Model	181
Figure 4.2 Study 2 Model	196
Figure A.1 Theoretical Development and Understanding	251
Figure B.1 Company Description.....	259
Figure B.2 Analytical Treatment.....	259
Figure B.3 Clout Treatment.....	260
Figure B.4 Authentic Treatment.....	260
Figure B.5 Negative Emotional Tone Treatment	261
Figure B.6 Negative Cascade Treatment.....	262
Figure B.7 Positive Cascade Treatment	263
Figure B.8 Essay 2: Study 2a Interaction Visualization.....	264
Figure B.9 Essay 2: Study 2b Interaction Visualization.....	264
Figure C.1 Overt Response Treatment	266
Figure C.2 Discreet Response Treatment.....	267
Figure C.3 Essay 3: Study 1 Interaction Visualization.....	267

Figure C.4 Analytical Treatment	268
Figure C.5 Clout Treatment.....	268
Figure C.6 Authentic Treatment.....	269
Figure C.7 Negative Emotional Tone Treatment	269
Figure C.8 Low Justification Interaction.....	270
Figure C.9 High Justification Interaction	270
Figure C.10 Full Interaction Visualization	271
Figure D.1 Consumer Inoculation Process	273

CHAPTER I

INTRODUCTION

Why do some consumers purposely create misleading information to share, manipulate impressions made online, and utilize online platforms to seek out and exploit others? Are those consumers potentially evil, or are they not so different, just using an online channel to build up their online image and create an online persona? Does malice drive them, or are they doing whatever it takes to make the best of opportunities? What impact do these individuals have on other consumers, and can a dynamic group impact a marketer's bottom-line? Market researchers have attempted to understand the dark aspects of consumer behavior through deviance (Moschis and Cox, 1989; Fullerton and Punj, 1997), misbehavior (Daunt and Harris, 2011), malicious behavior (Binns, 2012; Coles and West, 2016), and misinformation concerns (Larson and Denton, 2014), but little is known about the motivations of such misbehaviors. In today's marketplace, marketers can benefit from an understanding of the individual as well as the role of the online group in spreading misleading information. The essays presented in this dissertation are designed to provide a theoretical explanation of and solution to the damage caused to brands by this deviant consumer behavior.

Recent years have seen a surge in interest in dealing with the more dysfunctional aspects of consumer behavior (Payne and Frow, 2017). Much like in a psychology setting, this emphasis on the dark side of consumer behavior may be crucial in diagnosing ailments that may face consumer-brand relationships and providing optimal solutions. A literature search reveals that

more than 400 articles have been published in peer-reviewed marketing journals dealing with topics such as the dark side of “social media,” “CRM,” and “co-creation” since 2015. Prior and ongoing calls for studies by the *Journal of Promotion and Brand Management* (2018), *Journal of Marketing Management* (2019), and *European Management Journal* (2019) have explicitly devoted entire issues to addressing fake news, technological dystopia, and the dark side of social media.

With the proliferation of social media, consumers are gaining ground in social reach and can utilize online platforms to create and share misleading information when doing so helps to meet an end (such as financial, mischievousness, or power-seeking). For this reason, scholars have issued calls for research that provide in-depth theoretical and empirical perspectives into the drivers of dark behavior (Buckels, Trapnell, and Paulhus, 2014), anti-branding (Kucuk, 2010), and malicious behavior (Denegri-Knott, 2006). This dissertation, consisting of three separate essays, represents an attempt to answer those calls by addressing how misleading information is created, how it is disseminated, and how it can be eliminated. Each essay is described below in greater detail.

Essay One: Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information

Essay One employs qualitative and quantitative methods to explore both automatic and analytical cognitive components within autoinoculation by extending the lens of inoculation theory (McGuire, 1968) as a dual-process model of decision-making. The critical link between the Dark Triad (Machiavellianism, Narcissism, and Psychopathy) (Jonason and Webster, 2010; Jones and Paulhus, 2014) and misleading information sharing is established, the importance of both consumer proactivity and vigilantism (Saucier and Webster, 2010) is uncovered, and self-

justification (autoinoculation) of behaviors is developed to understand the reasons why consumers engage in such behavior.

A qualitative netnographic content analysis, examining ten Reddit forums, explores the thoughts and feelings associated with malicious online behavior, the hatred felt for brands, and the types of opportunities that are pursued. The emerging themes are then tested using structural equation modeling. The findings suggest that consumers' Dark Triad propensities affect both proactivity and vigilantism and thus strongly influence the likelihood to create and then share purposely misleading information with other consumers when the individual feels justified in doing so.

A key outcome of these studies is the finding that the Dark Triad, as a formative construct, is a predictor of a consumer's proactivity in seeking opportunities and utilizing online platforms to facilitate vigilantism. The propensities, when paired with cognitive justification for the action, influence the likelihood of a consumer creating and then sharing misleading information with others.

Essay Two: The Herd and the Bystander: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments

Essay Two explores the phenomenon of availability cascades in inducing groups of consumers to both accept and share misleading messages. As information becomes more available to others online, the original poster may be able to enact less effort for the message to spread, thus causing a fauxstorm, a firestorm of misinformation. This factor is essential for brands to understand since an active brand community may be able to take advantage of the same phenomenon to mitigate harmful information sharing. A better understanding of the components

that are under the brand's control to offset the dissemination of misleading information can benefit both researchers and practitioners.

While prior researchers have examined various brand community contexts, few have attempted to explain the impact these communities have in offsetting an anti-brand crisis. To address this deficiency in community research, inoculation theory as a dual-process model will be used as a theoretical lens to determine which message types, (Analytical, Clout, Authentic, or Emotional Tone), have the highest likelihood of acceptance by other consumer bystanders. The role of both Positive and Negative Cascades will be examined, challenges in creating counter messages will be explored, and the implications of the findings for researchers and practitioners will be discussed.

Essay Three: Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information

With the importance of the individual consumer established in Essay One, the role of the group found in Essay Two, the essay entitled "Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information" will examine more closely the role of halting the individual actions before the cascade has time to take effect. This essay examines the procedures that can be made early to influence malicious consumers before the critical mass of the availability cascade takes effect. Insights gained from Essay One regarding negative consumer behavior propensities allow an explanation of ways to find, seek out, and remove misleading information to avoid a potential brand crisis. A quarantine is introduced as a possible remedy for the misleading information by isolating the message before the Negative Cascade spreads the inoculation throughout the community.

A quantitative experimental analysis will provide insights that convey the best means of preventing the original poster from creating misleading information in the first place. The studies will determine when to offset the original message and what strategies should be implemented to do so.

Organization of the Dissertation

This dissertation is organized as follows:

Chapter Two contains the essay entitled “Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information.” This essay lays important groundwork establishing the importance of the Dark Triad and consumer propensities in influencing the creation and sharing of misleading information. Building upon that work, Chapter Three contains the essay entitled “Turning the Tide: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments.” This essay describes the impact that the initial message has on a larger group of consumer bystanders and the role of information availability in driving perceived accuracy in the message and subsequent sharing. Chapter Four contains the essay entitled “Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information.” This essay describes the role the original poster can play in offsetting a cascade of information and the ways to isolate the poster before the information is shared. These three essays are bridged through dual-process models of inoculation theory that explains the creation, dissemination, and elimination of misleading information. Finally, Chapter Five will synthesize the findings from the three essays; overall conclusions will be discussed, and a future research agenda will be established. These findings will provide essential insights to practitioners by establishing the role of inoculation in driving misleading information sharing, the role of the individual in initiating the process, the

role of the group in spreading the message, and the steps a company can take in combating this effect.

All research protocols described in this dissertation were submitted to the Mississippi State University Institutional Review Board for the protection of Human Subjects in Research. As part of this process, the primary researcher submitted examples of all recruiting materials, treatments, and surveys to IRB, along with descriptions of the respondent criteria and incentives for participation. IRB Approval was granted for each study.

CHAPTER II

ESSAY ONE

Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information

Abstract

Consumers who participate in online communication provide both benefits and challenges for brands. However, these individuals may pursue opportunities and coordinate communication with others when such online outlets offer a means to exploit both brands and other consumers. A consumer's Dark Triad propensities (Machiavellianism, Narcissism, and Psychopathy) can drive proactivity in seeking out and capitalizing on opportunities for personal enrichment. When posters find an opportunity, they can then act to sway and influence others. Consumers who experience vigilantism may believe that their view is right and must be shared with others. This study introduces a dual-process model of inoculation theory to the marketing and consumer literature by addressing the automatic components, through consumer propensities, and the autoinoculation that takes place when cognitive justification is made in sharing misleading information. Using a mixed-methods approach, this essay utilizes a netnographic content analysis with Linguistic Inquiry and Word Count (LIWC) to uncover qualitative insights shared in online forums and tests the relationships quantitatively through structural equation modeling.

“So, I just submitted a complaint to customer service about their stonewalling, feedback manipulation, and basic dishonesty. Please take screen shots of their feedback page at various times in order to have proof of 1-star reviews being removed. I'm doing the same on my end.”¹

Introduction

In the age of social media, consumers can connect quickly and easily to many other consumers to share both positive and negative experiences. Studies have suggested that this means of communication could be perceived as more accurate and believable since customer-generated information can seem more authentic than company-sponsored marketing campaigns (Render, 2018; Kwok and Yu, 2013; Pantelidis, 2010). However, such forums can be used to share resentment or displeasure regardless of situational context or evidence. If these platforms go unchecked, companies can quickly lose control of online conversations. What is more disconcerting is the ability for posters to create misleading information to share with others (Dubois, Rucker, and Tormala, 2011; Visentin, Pizzi, and Pichierri, 2019).

Even social media giant Facebook recently came under scrutiny for not providing checks and balances for information that is shared on the platform. Because posters have found loopholes in monitoring, purposely misleading information can be created and shared with others with little effort. Moreover, companies have voiced concern regarding false and deceptive information sharing since the messages can damage both the brand's social and fiscal bottom lines (Loeb, 2019). Although firestorms of negative information can plague a brand after a

¹ *An online consumer on Reddit was vigilantly describing his reaction to the perceived unethical actions of an online retailer after a pricing mistake. The poster acted to voice the view that the reviews should be monitored and coordinated.*

failure, damaging information may come from the tactical planning of one individual in misleading information sharing. Both companies and consumers are becoming concerned with the presence of post-truth politics (Roberts, 2016), truthiness in information sharing that “sounds” accurate (Meddaugh, 2010; Munger, 2008), and fake information sharing aimed to cause harm (Newman, Garry, Bernstein, Kantner, and Lindsay, 2012). Trolling behavior, acting in a destructive or disruptive manner in an online setting with no apparent purpose, can have devastating consequences for brands. Examples of companies on the receiving end of this behavior include an online retailer encountering misleading information that was spread to other consumers about a pricing mistake and Monsanto, who was attacked by online trolls to influence the local production of agricultural products. Consumers online also engage in coordinated misleading information through negative reviews on Yelp, TripAdvisor, Amazon, and IMDB to cause intentional harm to brands and companies. This action is particularly disconcerting since 75% of business owners state that online reviews are important, and one-sixth of those owners say bad reviews could ruin their business (Pickard-Whitehead, 2017).

Misleading information sharing differs from traditional word-of-mouth after service failures in that the said event expressed may not actually occur, and the created information is intended to deceive other consumers. Possible drivers of this malicious consumer behavior are a set of personality characteristics referred to as the Dark Triad traits. The Dark Triad is defined as the proclivity an individual has toward cunning and manipulation (Machiavellianism), self-grandiosity and sense of importance (Narcissism), and callousness and impulsivity (Psychopathy) that can be exerted towards others (Jonason and Webster, 2010; Karampournioti, Hennigs, and Wiedmann, 2018). Although psychopathic (Parry, 2011) and narcissistic (Peisley, 2017) diagnosed disorders affect a rather small portion of the overall population, with each

representing about 1% of the general population, otherwise healthy individuals can also exhibit these tendencies. For instance, a manager may act to manipulate others to meet a deadline, a performer may feel an inflated sense of ability, and a teacher can act somewhat callously toward students. Therefore, all people exhibit these propensities, with higher levels relating to mischievous and malicious behavior (Jonason and Webster, 2010).

The Dark Triad impacts consumer complaining, trolling, brand and employee bullying, and lying intentions (Breitsohl, Roschk, and Feyertag, 2018; Harrison, Summers, and Mennecke, 2016). The Dark Triad has also been demonstrated to have a direct positive relationship with deviant (unwanted and non-normal) misbehavior intentions (Daunt and Harris, 2011), and exploitative tendencies (Jonason, Webster, and Schmitt, 2009), but also with innovation (Wisse, Barelds, and Rietzschel, 2015). These findings suggest that individuals with higher Dark Triad tendencies are also skilled at proactively seeking out and exploiting opportunities. This proactivity in seeking opportunities is defined as a proclivity that individuals have to pursue actions that influence their current state or social environments (Bateman and Crant, 1993; Seibert, Kraimer, and Crant, 2001)

Research further suggests that consumers who possess this penchant for proactivity actively seek to influence others (Bjorklund, Bhatli, and Laakso, 2013; Pitt and Ewing, 2002). This influence may stem from a desire to share one's views and coordinate the actions of others for one's own agenda. Vigilantism illuminates the belief that an individual has in which one's views are superior, and therefore, the individual must spread these beliefs to others (Saucier and Webster, 2010). The majority of research on consumer information sharing thus far has focused on an initial failure driving information sharing within consumer communities (c.f. Larson and Denton, 2014; Daunt and Harris, 2011; Kucuk, 2010; Herhausen, Ludwig, Grewal, Wulf, and

Schoegel, 2019). However, sometimes the posters share misleading online information “just because,” without any apparent motivation. To date, no published research examines the role of the Dark Triad, proactivity, and vigilantism in driving intentionally misleading information. Since posters online can behave maliciously in real-time, develop forums for a specific agenda, and can easily engage others, this topic is a subject worthy of research to understand these individuals and the influence they can exert on others (Rauschnabel, Kammerlander, and Ivens, 2016), message types that are most effective (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019), and content that reinforces a message (Hewett, Rand, Rust, and van Heerde, 2016).

The studies that follow contribute to the understanding of how posters actively persuade not only others but also themselves. Inoculation theory will be used as a theoretical lens and states that individuals can be immunized from counter-information by introducing others to a weak counterargument that can be easily dismissed (Banas and Rains, 2010). Just as balanced advertising appeals have been shown to reduce counterarguing (c.f. Hunt, 1973; Kim, 2013; Veil and Kent, 2008; Eisend, 2006), so too does information inoculation serve to prepare an audience to battle contrary information when encountered. For example, a brand can enact a campaign after a crisis to thwart the efforts of competitors trying to exploit it. A study by Mikolon, Quaiser, and Wieseke (2015) found that customers who were inoculated toward perceptions of satisfaction responded less harshly after a failure by reinforcing preconceived heuristics of prior satisfaction. In the case of misleading information generated by the consumer, it appears they are implementing “marketing tricks” of their own. Necessarily, the original poster can act to inoculate himself (autoinoculation through automatic and analytical cognitive resources) and then inoculate others within an online social environment.

While researchers have addressed the roles of online communities and consumer responses to failures, little research has examined the role of dark individual traits as a holistic driver of creating and then posting misleading information that potentially initiates a brand crisis regardless of a failure occurring. Additionally, very little information is available to describe the types of opportunities that consumers influence the likelihood to create and share misleading information to cause harm. This research addresses that paucity by combining both qualitative methods (a netnographic content analysis with a Linguistic Inquiry and Word Count supplement) and quantitative methods (structural equation modeling) to explore the creation and dissemination of purposely misleading information with consumers online and to determine both the motivations and individual traits that are present in these consumers. Essay One addresses the traits that drive this behavior through both automatic means and analytical means.

The paper is organized as follows: First, the literature regarding negative online information sharing and deviant online behavior is described. Then the methods are explained, followed by a description and discussion of the qualitative results. Finally, the qualitative conclusions are triangulated by quantitatively applying the emergent theory from the analysis, implications and findings are discussed, and future research directions are recommended.

Literature Review

A rich body of literature exists that explains both positive and negative consumer actions within the marketplace through advocacy and deviance. When consumers advocate for a brand offering, the consumer acts to protect beloved brands (Wallace, Buil, and Chernatony, 2014). Additionally, loyal customers may serve as a consumer vigilante who acts to mitigate harm to the brand (Muniz and Schau, 2007). However, these advocacy actions can provide negative outcomes for a brand when the consumer begins to defect from the traditional buyer/seller

paradigm set by the marketplace. For instance, consumers may act as active advocates for other consumers and may monitor price fairness and positive word-of-mouth by acting out against a brand (Wakefield and Inman, 1993; Larson and Denton, 2014). These consumers may also begin to monitor a brand's marketing communications diligently to ensure that accurate or fair information is shared by organizations (Hsiao, Shen, and Chao, 2015).

Overall, consumer and brand advocacy groups aim to protect consumers from misleading tactics or being “duped” by marketing campaigns (Holt, 2002). The outcomes of such programs can contribute to discussions made in consumption communities that compare alternatives (Stokburger-Sauer and Wiertz, 2015) and result in intentions to avoid brands or offerings that do not meet the needs or expectations of a community (Jayasimha Chaudhary, and Chauhan, 2017). This advocacy for other consumers can influence the negative behaviors and information directed at brands. When negative motivations act in tandem with personal justification for a consumer's behavior, undesirable consumer actions may result.

While some level of consumer advocacy can lead to positive sharing of unsponsored communication, consumers can engage in practices that may actively harm an organization. For instance, the concept of consumer deviance encompasses both fraudulent and negligent behaviors (Moschis and Cox, 1989). Fraudulent deviance typically describes consumer behaviors that are explicitly illegal and cause harm to the firm. These behaviors can include shoplifting, causing physical damage, and other practices that are overtly illicit (Dootson, Lings, Beatson, and Johnston, 2017).

Although fraudulent deviance must be monitored, and typically is, negligent deviance or actions that consumers take that deviate negatively from the expectations of the firm, may be equally important to track and manage. Whereas overtly fraudulent behavior may be more

visible, negligent behaviors are even more insidious because they can take place right “under the nose” of the brand. For instance, just like fraudulent, negligent behaviors such as excessive returns, complaints, and harmful information sharing can negatively impact overall profitability (Fullerton and Punj, 1997; Daunt and Harris, 2011). Consumer deviance can also extend beyond in-store actions to online discussions and information sharing (Binns, 2012; Coles and West, 2016).

These deviant behaviors must be addressed to ensure that proper relationships are maintained (Daunt and Harris, 2011). Consumers can actively engage in deviant online behavior to exploit price differences, share negative reviews, and organize group responses to confront a brand (Denegri-Knott, 2006). Moreover, when consumers with ulterior motives to cause harm to others begin to share maliciously misleading information about a brand, this consumer deviance can become even more problematic once both websites and social media discussions are created.

Prior literature has addressed the development of anti-brand websites (c.f. Kucuk, 2008; Kucuk, 2010). These websites are typically developed to convey a level of resentment, distaste, or hatred toward a focal brand. Kucuk (2008) provided examples such as “starbucked.com,” “aolsucks.com,” “Mcspotlight.com,” and “Walmart-blows.com,” which are created by consumers to vocalize discontent. Many consumers may accept negative stories shared by others or personal views as fact since the receivers tend to piece together a narrative based on stories that are told by others (Render, 2018; Pickard-Whitehead, 2017; Kahneman, 2011). Consumers may act vigilantly by questioning company responses, information communicated, and prior experience from similar situations to ensure that the information posted online or in online reviews is accurate (Larson and Denton, 2014). However, these websites are somewhat static since viewers must specifically seek them out and require financial resources for upkeep.

Because these sites are less malleable and do not typically evolve or adapt to new information from multiple sources, the websites may not contribute to further online discourse beyond their focal viewers. Therefore, companies may be able to monitor these websites more efficiently.

In today's age of social media, consumers have a much greater ability to leverage more substantial scale effects on social media use. For example, a small computer company, rather than Nissan Motors, owns "Nissan.com." The organization voices ongoing legal battles with the automaker and has developed a forum for consumers to voice disdain for Nissan motors (Nissan.com). When consumers have this kind of access to others, some individuals exhibit vigilante characteristics through conveying consumer advocacy motivations and attempt to protect others from companies that are deemed unethical or immoral from the poster's perspective (Kravetz, 2007; Coles and West, 2016). However, when this vigilantism is prompted by a self-justified need to share opinions with others, without regard for accuracy, potential damage can be done (Saucier and Webster, 2010). For example, a consumer that experiences some level of bias against a brand can develop a propensity toward consumer social vigilantism. This propensity is marked by the view that the consumer's individual opinion is the only accurate view, and it is their responsibility to "enlighten" others at all costs rather than provide evangelistic support for a brand or group (Saucier, Webster, Hoffman, and Strain, 2014).

Additionally, online forums, such as those available on Reddit, can be created in real-time (while an event is occurring) to develop a multiplier effect by engaging others to a coordinated end (i.e., coordinated information sharing, purchasing, or comparisons). These effects also take place in a somewhat anonymous online environment that contributes to group-level authenticity through collaboration rather than relying on the credibility of one source (Rains, 2007). In spite of these volatile conditions, little research exists that determines the

motivations, traits, and outcomes of such consumer-coordinated efforts in online communities without the brand's control. The information presented in this literature review that took place *a priori* informs the qualitative investigation that follows.

Sensitizing concepts such as the Dark Triad traits (Machiavellianism, Narcissism, and Psychopathy), consumer motivations (whether monetary, ego-driven, or pleasure), and potential outlets for information sharing served to ground the netnographic content analysis in Study 1. For three years, the researcher explored the related topics in numerous papers and readings (see APPENDIX). Other issues that were considered before conducting the netnographic content analysis included concepts of equity, justice, and attribution. The concept of the Dark Triad was identified as a possible explanation for deviant consumer behavior since individuals may act maliciously and mislead others without any real prompt or provocation from a brand (i.e., failures, negative advertisements, or encounters) and prior studies have linked the Dark Triad to trolling behaviors (Buckels et al., 2014).

Motivations are also of interest because they can work in tandem with negative personality traits. Consumers who feel that they can further their wellbeing by influencing others may also do so by using less than desirable means. Therefore, a group atmosphere within an online community bolsters the motivations since they group dynamic can be used to reinforce views and provide outlets for information dissemination (Hewett, Rand, Rust, and van Heerde, 2016; Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). The group setting of the online platform facilitates synchronous discussions and proactive information-sharing with others to convey a particular point or belief. Therefore, these concepts will drive the netnographic content analysis by providing an initial grounding in the literature.

This qualitative study served to identify organic themes that emerged within the online discussions. Retrieving organic themes ensures that the literature review guides the study rather than only confirming prior expectations. By addressing the key drivers of decision-making, through both heuristics (mental short cuts) and analysis (greater cognitive resources), in-depth insights can be retrieved through a more pluralistic understanding through numerous online interactions and perspectives (Gummesson, 2005). Moreover, the qualitative insights will act to develop a more compelling argument for the contribution of the research and the quality of the quantitative analyses (Richardson and St. Pierre, 2008).

Qualitative Research

Study 1: Netnographic content analysis

Just as rich insights can be gained from ethnographic research that allows a researcher to observe the behaviors of a group of people (Elliot and Jankel-Elliot, 2003), so too, can consumer behavior be observed in an online or digital setting (Kozinets, 2002). A netnography can be implemented to understand the interactions that take place between consumers in brand communities (Brown, Kozinets, and Sherry, 2003), develop cultural insights (Kozinets, 2002), analyze online product reviews (Yang and Peterson, 2003), and create a means to gain unobtrusive insights into sensitive topics (Langer and Beckman, 2005). Therefore, an observational netnographic content analysis provides an appropriate avenue in determining the motivations and traits of those engaging in malicious or misleading online behavior without interrupting the environment itself.

To determine the traits and core motivations for consumer sharing of misleading information and malicious behavior online, an organic understanding of the nuanced relationship of the poster and the online environment must be studied. The online disinhibition effect

describes a phenomenon in which individuals may take on a different online persona that is less constrained than their typical offline personas (Suler, 2004). This effect can take place since individuals can capitalize on the anonymity perceived in an online environment or create a disconnect from everyday life. Therefore, observing the context and audience of online communication can be crucial to understanding the behavior as online communities can develop close-knit and open communication sharing between members (Muniz and O'Guinn, 2001; Algesheimer, Dholakia, and Herrmann, 2005).

The purpose of Study 1 is to uncover and understand the negative consumer propensities and motivations that drive the creation and dissemination of misleading information. The themes that emerge will then be subjected to quantitative testing. This allows the researcher to triangulate the findings, thus enhancing their generalizability. In the following sections, the research design, literature review, sample descriptions, analysis, and key findings are presented

Netnographic Content Analysis Research Design

Qualitative research can be implemented when crucial insights may be challenging to assess initially in a quantitative context or when the core drivers of the relationship are not specifically understood (Corbin and Strauss, 2008; Mittal, 2015). A netnography is a virtual form of a traditional ethnography that allows a researcher to explore consumer motivations, beliefs, and traits in a natural online environment rich with data (Kozinets, 1997; Brown, Kozinets, and Sherry, 2003). A common ecosystem for online information sharing is found on Reddit. Reddit is a social media site that aims to provide users with a way to aggregate news, rate web content, and coordinate discussions (Redditinc.com). Since the sole purpose of Reddit is to not only create but also share and rate content throughout a rich online ecosystem, it is specifically chosen as an appropriate sample over other platforms (i.e., Twitter, Facebook, Instagram) to represent

how topics are shared, how conversations are facilitated via user to user, and how user acceptance determines the visibility of messages.

Subreddits can be created by individual users to generate group discussion on topics. Since the focal participants in the subreddit discussion forums operate in close-knit communities, an observational symbolic netnography (non-intrusive) is implemented to observe the ongoing discussions and behaviors without interacting in the organic environment (c.f. Kozinets, 2002; Kozinets, 2015; Langer and Beckman, 2005). As data are publicly accessible online through the subreddit forums, data can be readily retrieved without disrupting the consumer interactions (Nelson and Otnes, 2005). Once the netnographic content analysis is completed, a conceptual model can be developed to address the role of the themes in driving the sharing of misleading information with others.

This netnographic content analysis research explores online consumer deviance and information sharing vigilance through an examination of the traits, motivations, and behaviors of online subreddit participation. Therefore, the netnographic content analysis addresses the following question: “What drives individuals to create and then share misleading information?” First, Google was used to find discussions of malicious information sharing, trolling, anti-brand, and anti-company conversations on internet forums and within social media. Popular press articles, academic articles, and numerous forums that convey discontent were identified as potential sources of information.

Kozinets (2002, p.63) suggests that “online communities should consist of 1) a relevant group of participants, 2) higher “traffic” of posters, 3) large numbers of posters, 4) descriptively rich data, and 5) adequate “between poster” communication.” This netnographic content analysis covers C2C communication on numerous subreddits. Subreddits are created and moderated by

individual Reddit users to explore and discuss specific topics with others. Reddit communities are often quite active and are used in the analysis to form categories based on the motivations and traits of participants. Reddit is deemed an appropriate sampling frame since the site is completely user-driven, and new topics can be created and discussed promptly. Additionally, Reddit forums have a large online footprint since prior topics can be searched after long after the discussion has ended.

This netnographic content analysis focuses on participant interactions observed by the researcher through subreddits such as “why do people troll on the internet” (228 comments with 103 participants), “what psychological reasons do you have to troll” (128 comments with 55 respondents), “companies likely to scam customers” (2,100 comments with 951 participants), “corporate greed examples” (3,800 comments with 145 participants), “hated brands (2,000 comments with 85 participants), and “Reddit Ask Me Anything –fake complaints and reviews” (331 comments amongst 23 participants). Topics are then applied to an additional subreddit forum focused on consumer actions that follow a pricing mistake.

The themes uncovered through the netnographic content analysis are reflective of the conversations that take place between participants across ten different online subreddit forums to determine the motivations and traits that are discussed. Online community member anonymity is maintained throughout as Reddit users employ pseudonyms, a practice that likely enhances the online disinhibition effect and provides a view that is likely less affected by social desirability bias. Companies and brands that are discussed are disidentified to ensure that they cannot be identified. Because the observational symbolic netnographic content analysis does not allow the researcher to directly question the informants regarding their motivations to post within these forums, an additional tool is used to triangulate the researcher’s observations. Specifically,

LIWC (Linguistic Inquiry and Word Count) is conducted on the comments to allow for additional insights into the online discussions.

LIWC

LIWC software allows a researcher to select a specific written text type and then compare the selected text to qualitative and quantitative baselines that have been established through numerous algorithms and analyses by expert judges (Pennebaker, 2011). Additionally, LIWC's dictionary is updated regularly and now consists of over 5000 words (Pennebaker, 2019). The software has been implemented to address marketing phenomena such as online feedback loops (Hewett, Rand, Rust, and van Heerde, 2016), viral marketing (Berger and Milkman, 2012), online reviews (Ludwig, De Ruyter, Friedman, Brüggem, Wetzels, and Pfann, 2013), negative affect in information sharing (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019) and audience size (Barasch and Berger, 2014). This research focuses on messages classified by LIWC as "Social media: Twitter, Facebook, and Blog" within the LIWC software to analyze the forum comments. The netnographic content analysis focuses on the nuances associated with the individuals participating within the forums, and LIWC helps to provide additional context to the posters themselves.

The LIWC output identifies the percentage of the words in the message that contain I-words (I, Me, My), social words (words dealing with groups), positive emotions (happy, content, excited), negative emotions (angry, upset, worried), and cognitive processes (words that imply cognition such as causation, discrepancy, tentativeness, and insights). LIWC does not remove the researcher's judgment from the analysis but acts as a supplement to the netnographic content analysis by providing additional insights and confirmation of researcher insights. This coding

software adds a quantitative element to the analysis by providing supplemental context to the data.

LIWC also provides the following summary variables: analytical thinking, clout, authenticity, and emotional tone. Each summary variable is explained below. The LIWC website indicates that these variables are based on previous language research and are developed with an iterative coding process (Mehl, 2006; Kacewicz, Pennebaker, Davis, Jeon, and Graesser, 2013; Newman, Pennebaker, Berry, and Richards). The variables are based on standardized scores ranging from 0 to 100 (converted to percentiles). The complete LIWC summaries for selected comments are reported within the netnographic content analysis results section. An example LIWC output is posted below in Figure 2.1 to aid in the discussion of the summary outputs.

LIWC Example	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog	
	“I went for an interview recently at a prestigious healthcare organisation. Felt pretty proud of myself for making it through the first 2 rounds. I had a good feeling I’d get in since it was the final round. But NO. Throughout the entire interview, the panel of 4 mocked me, said I wasn’t good enough, and said I didn’t know to shake hands before the interview. WHEN ALL OF THEM SAT IN THEIR COMFY CHAIRS AS I APOLOGISED AND SHOOK EACH INDIVIDUAL HAND.”	I-words (I, Me, My)	9.6	5.51	
		Social Words	12.0	9.71	
		Positive Emotions	4.8	4.57	
		Negative Emotions	3.6	2.1	
		Cognitive Processes	13.3	10.77	
		Summary Variables			
		Analytic	86.9	55.92	
		Clout	31.5	55.45	
		Authenticity	97.2	55.66	
		Emotional Tone	48.1	63.35	

Figure 2.1 LIWC Summary Outputs Example

The analytical thinking variable is a factor-analytically derived dimension based on eight cognitive function word dimensions (c.f. Pennebaker, Chung, Frazee, Lavergne, and Beaver 2014). The analytic dimension addresses the degree to which people use words that suggest formal, logical, and hierarchical thinking patterns (Pennebaker, 2019; Pennebaker, Chung,

Fraze, Lavergne, and Beaver, 2014). People low in analytical thinking tend to use language in more narrative ways, focusing on personal experiences. This dimension helps determine the level of effort and thought in generating the discussion posts.

The clout variable refers to the relative social status, confidence, or leadership that people display through their writing (Pennebaker, 2019). The algorithm was developed based on the results from studies where people interact with one another (Kacwicz, Pennebaker, Davis, Jeon, and Graesser, 2013). Individuals may convey a certain level of importance or status based on their postings. This variable provides insights into the poster's perceived sway or ability to influence others.

Posts that appear to be authentic can reinforce the message that is being conveyed. Newman, Pennebaker, Berry, and Richards (2003) developed the algorithm for authenticity through a number of studies in which people were convinced to be honest or deceptive. Additionally, Pennebaker (2011) contributed to the creation of the algorithm through published summaries of deception studies. When people reveal themselves authentically or honestly, they are perceived as more personal, humble, and vulnerable (Pennebaker, 2011). Additionally, posters may convey information in a way that is deemed more authentic to influence the other members of the online group.

Finally, LIWC provides a variable for the emotional tone of comments through both the positive and negative emotion dimensions (Cohn, Mehl, and Pennebaker, 2004). Pennebaker (2011) states that the algorithm is built so that the higher the number (i.e., 100), the more positive the tone. Moreover, the authors suggest that scores below (50) indicate a more negative emotional tone (Cohn, Mehl, and Pennebaker, 2004). This measure can begin to address the raw emotion that is conveyed with the forum posting.

Findings Related to Online Information Sharing and Trolling Behavior

Motivations for Sharing Misleading Information

The netnographic content analysis first focuses on uncovering the motivations of those engaging in fraudulent and misleading information sharing online. Buckels, Trapnell, and Paulhus (2014) discuss the concept of online trolling behavior as activities aiming to solicit some responses from others. Trolling behaviors are distinct from online bullying in that trolling behavior typically relies on deceptive practices and perceived randomness (Buckels, Trapnell, and Paulhus, 2014; Binns, 2012). Although large companies are no stranger to large-scale attempts to misinform or mislead both employees and customers, small-to-medium sized companies can also become targets and are becoming concerned with online trolls (Pickard-Whitehead, 2017; Meschke, 2014). The netnographic content analysis reviews posts by posters who either self-identify as “trolls” or interact with those who identify as “trolls.” Existing message boards created by Reddit users to discuss “why do people engage in trolling?” and “what are the psychological reasons for trolling?” are examined to answer the questions posed in this study.

The subreddit focused on the actual trolling behavior consisted of 228 comments amongst 103 participants, and the subreddit asking those involved in the behavior to reflect on the reasons for trolling included 128 comments and 55 participants. The need for power and control quickly emerged as a theme discussed by the participants. For example, posters posted messages such as the following to suggest the power motivations of those engaging in online trolling behavior.

“Because the very base level of what trolling is, is the means to convince yourself you've got power over someone else”

“There are people out there who feel completely ineffectual in the real world and they find it thrilling and intoxicatingly powerful that they can actually have some kind of effect online”

The online forums cited academic research such as studies on the online disinhibition effect (i.e., Suler, 2004) and studies on trolling behavior (i.e., Buckels, Trapnell, and Paulhus, 2014). Specifically, posters mentioned online anonymity leading to the creation of a persona that is different from their real identity, and that individuals may say or do things they would not do in an offline setting (Denegri-Knott, 2006). Such discussions within the forum indicate that the conversations are not only insightful but also thoughtful in determining an understanding of the behavior. One poster stated the reason for the behavior is:

“Freedom from real consequences, mainly. It's human nature for young people and damaged people to lash out and behave badly if there are no consequences”

and another stated:

“It's because it's funny. That's literally it.”

The posters in the forum engaged in introspection to develop and articulate meaningful explanations for both current and past behavior. They opened up to others in describing the actions and the harm that could come to others. Additionally, posters began to solidify points of view by providing specific details of prior behavior and sharing information aimed at meaningfully and effectively answering the questions posed by the subreddit forum.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
1	“It's the illusion of power, very addictive little dopamine hit/feedback loop. The less powerful/significant you are in real life, the greater the rush of power you feel when making people upset or angry on the internet.”	I-words (I, Me, My)	0	5.51
		Social Words	7.9	9.71
		Positive Emotions	2.6	4.57
		Negative Emotions	5.3	2.1
		Cognitive Processes	7.9	10.77
		Summary Variables		
		Analytic	92.8	55.92
		Clout	85.4	55.45
		Authenticity	33.4	55.66
Emotional Tone	2.5	63.35		

Figure 2.1 Poster 1 Summary

The LIWC standardized scores output for Poster 1 above indicated that the message was in the 90th percentile in conveying an analytical statement and within the 85th percentile in the clout variable. Therefore, the poster is providing evidence of formal thinking patterns while expressing confidence and knowledge of the topic. The informants in the netnographic content analysis are compared to the baseline created by Pennebaker (2011) to determine how the posters in the Reddit forums compare to the average social media poster. Moreover, many participants stated that an illusion of power influenced the pursuit of the behavior. For instance, the LIWC output for Poster 2 indicates that the individual below focused on his own experience and utilized (11.5%) “I-words” compared to the average (5.51) for average blog posts. The commenter reflected on personal experiences in which they engaged or how they were affected by similar behavior. The second poster’s message consisted of (15.4%) in cognitive process words, greater than the average (10.77%) for blog posts. The introspective focus helped to build the analytic nature of the comment, (73rd percentile), and the conveyed authenticity in the

message, (99th percentile). The LIWC output indicates that Poster 2 is reflecting analytically on actions of his/her youth and is conveying the information honestly or humbly.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
2	“I trolled in my younger years. The idea that 13 year old me could anger someone across the world with no consequences was a power thing.”	I-words (I, Me, My)	11.5	5.51
		Social Words	3.8	9.71
		Positive Emotions	0	4.57
		Negative Emotions	3.8	2.1
		Cognitive Processes	15.4	10.77
		Summary Variables		
		Analytic	73.9	55.92
		Clout	12.4	55.45
		Authenticity	99	55.66
		Emotional Tone	1	63.35

Figure 2.2 Poster 2 Summary

However, the participation in misleading online behavior extended beyond the core need for power by addressing the use of online communication sharing as a means to manipulate others and fulfill a specific need/agenda. This opportunistic motivation suggests that those who engage in online malicious behavior may not focus solely on the power aspect, but more so on proactivity toward pursuing opportunities that fulfill a given need. These posters suggest that people may need power but also may utilize trolling behavior to manipulate or influence others for their own gains. These needs may focus on placing actual emotional strain on others, creating harm to others, or the exploitation of others. For example, Poster 3 states that people that purposely try to mislead or harm others may feel that they have some control over the receivers of the information and can influence the sought-after outcome.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
3	<p>“They feel validated when you are affected by their actions. They feel like they have the power to to [sic] manipulate you, they confirm in their heads that they are better than you. They won't care about what you'd say about them, but you obviously care about what they have to say. They have the power. You don't.”</p>	I-words (I, Me, My)	0	5.51
		Social Words	33.9	9.71
		Positive Emotions	5.4	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	12.5	10.77
		Summary Variables		
		Analytic	4.3	55.92
		Clout	99	55.45
		Authenticity	7.8	55.66
		Emotional Tone	97.9	63.35

Figure 2.3 Poster 3 Summary

The LIWC output for this poster conveys the perceived motivations of malicious online behavior and warns others through the use of social clout (99th percentile) and emotional tone (97th percentile). The individual is conveying a sense of status in describing the behavior while also expressing a positive emotional tone in doing so. This finding may indicate that the individual could feel good that he or she is sharing perspectives, albeit negative information. Additionally, Poster 3 uses words to describe groups and cognitive thought. Suggesting that the activity might stem from a broader set of motivations, beyond a sole focus on power online, that pursues forms of manipulation to seek out and exploit opportunities by influencing others.

When these activities extend beyond individuals to brands/companies, online behavior can then influence other consumers. This influence can then be used to reinforce negative information. When a bias arises, participants may utilize additional cognitive resources to justify their bias (Kahneman, 2011). Additionally, the forums referenced the concept of “schadenfreude” (i.e., pleasure in someone else’s misfortune) as a core motivator of the online malicious behavior.

“I can say I've enjoyed trolling in games in the past and it really is that it feels good to watch someone else suffer”

Additionally, the key driver for a negative review could be to not only incite others but also enjoy watching the other party respond. Those that post negative information tend to view the focus of the information as deserving of the treatment (Chester and DeWall, 2017; O'Meara, Davies, and Hammond, 2011).

“Sit back and watch for pleasure”

“I think people enjoy seeing opinions they disagree with shut down harshly. It's very satisfying.”

This drive, seemingly influenced by vigilantism, the belief that an individual's views are superior and must be spread to others, in some cases for a “greater good,” became especially apparent in online Reddit forums. These forums discussed “companies likely to scam customers” (2,100 comments with 951 participants), “corporate greed examples” (3,800 comments with 145 participants), “hated brands” (2,000 comments 85 participants), and “Reddit Ask Me Anything – fake complaints and reviews” (331 comments amongst 23 participants).

“ I would think so, but my once a month or so complaints shouldn't damage them financially too much”

“Fake reviews are more likely to tell a story or try too hard to justify why you were there.”

Opportunistic actions by consumers can be justified based on the perception of the company or the focal brand. For instance, a poster in an “Ask Me Anything” subreddit stated that a larger organization would not notice the financial impact of false complaints. Therefore, multiple locations could be targeted for exploitation over time. Since Poster 4 utilizes the

subreddit to convey his/her opinion and attempts to sway the opinions of others, he shows signs of vigilantism. The poster exhibited levels of vigilantism in not only the actions (of taking the matters into his own hands) but also the descriptions of the activities to other posters (by encouraging step-by-step replication by the other posters).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
4	<p>“The easiest method is to simply email the company of your choosing. For this example ill use McDonald’s. “Say something that cant be easily verified The bread was stale.” “The meat was overcooked.””....I got 4 “free value meal” coupons.</p>	I-words (I, Me, My)	2.5	5.51
		Social Words	10	9.71
		Positive Emotions	10	4.57
		Negative Emotions	2.5	2.1
		Cognitive Processes	10	10.77
		Summary Variables		
		Analytic	77.3	55.92
		Clout	69.1	55.45
		Authenticity	1.3	55.66
		Emotional Tone	99	63.35

Figure 2.4 Poster 4 Summary

Not only did Poster 4 state how s/he has purposely misled a company but also provided additional steps on how others could also engage in the behavior. Since vigilantism is prompted by a need to share opinions with others (Saucier and Webster, 2010), this poster attempts to reach a larger number of “others” as seen in the forum discussing corporate greed examples. For instance, this was not uncommon as 67% of the posts were created by those actively sharing opinions, experiences, and information, while the remaining 33% responded to the original posters. The LIWC output indicates that Poster 4 portrayed him/herself as a body of knowledge using analytical communication (77th percentile) and clout (69th percentile).

Moreover, the poster seemed to take pleasure in using the platform to voice the opportunistic behavior with a high emotional tone (99th percentile).

By reviewing the forums, negative information sharing seems to be seen as appropriate when companies are viewed in a negative light. Negative views could be due to perceptions of employment practices (21%), negative personal experiences (25%), environmental practices (4%), customers they attract (16%), or company reputation (34%). When participants begin to address similar themes, the messages can start to converge on specific companies. Additionally, the availability of information from a bystander’s perspective can influence the perceived accuracy of a statement, whether it is a fact or fiction. This role of availability is evident within a Reddit “Ask me Anything” revolving around a person who is incentivized to provide fake reviews on behalf of companies.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
5	“I’m a writer for my main job and I browse the Craigslist writing gigs section for ghostwriting and odd jobs I can do remotely since I’m a fast writer and looking for some extra cash. An agency posted an ad and I was applying for everything at the time. They emailed me back after reading some of my publications and we set up a contract. The pay is shit, but it’s kind of fun and it’s just a side job so I can quit anytime I want with no consequences. I get sent projects to complete in 24 hours so I either do it at work or when I get home. Usually at work if I’m being honest.”	I-words (I, Me, My)	12	5.51
		Social Words	4.3	9.71
		Positive Emotions	1.7	4.57
		Negative Emotions	0.9	2.1
		Cognitive Processes	12.8	10.77
		Summary Variables		
		Analytic	40.3	55.92
		Clout	11.6	55.45
		Authenticity	88.5	55.66
		Emotional Tone	41.2	63.35

Figure 2.5 Poster 5 Summary

Poster 5 states that the more similar the fake reviews are to the authentic reviews, the more likely others are to perceive them as genuine. The poster explicitly states that the reason for the fake reviews is the pursuit of a financial opportunity, but that this pursuit is further encouraged by the opportunity to share the information with other parties but also to educate posters about the

creation process. This poster laboriously created a message based on cognitive processes (12.8%) and high levels of conveyed authenticity in the message (88th percentile).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
5	“When I'm reviewing anything, I read previous reviews to get a feel of what people like and don't like about the place or thing.”	I-words (I, Me, My)	8.3	5.51
		Social Words	4.2	9.71
		Positive Emotions	4.2	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	12.5	10.77
		Summary Variables		
		Analytic	26.8	55.92
		Clout	10.6	55.45
		Authenticity	91.3	55.66
		Emotional Tone	92.4	63.35

Figure 2.5 (continued)

Moreover, the poster's emotional tone score (92nd percentile) indicates that the individual is communicating positively when describing the opportunistic venture, although the actions are viewed in a negative light by others. For instance, Poster 5 exhibited a high positive emotional tone in the presences of other posts, such as the following:

“It's depressing that people feel the need to do unethical (or immoral) things to make money. It completely ruins any sense of trust in society. I don't tend to judge but I wouldn't put "integrity" on the top of the list of your traits.”

Therefore, individuals can employ cognitive resources to justify the actions that are taken and allow for the positive emotional resonance in information sharing. This finding indicates that people that are proactively pursuing opportunities can reinforce and justify their views and actions from the contrary views of others regarding the practice.

In analyzing the forums, the key themes that emerged revolve around consumer proactivity toward pursuing opportunities, whether it be through power-seeking (19%), amusement (24%), entertainment (14%), financial gain (7%), or mischievous enjoyment (36%). When consumers develop this disinhibition from their “in-person” persona and disconnect in an online setting, they may use forums and message boards to convey experience and expertise linked to a brand, company, or product regardless of facts. As these individuals capitalize on their ability to share information with others quickly, the message can promptly diffuse throughout the online discourse. Additionally, the posters aim to justify their desire to create information before sharing it with others in an online setting. The more often these are viewed or reposted, the greater the chance that they are perceived as factual messages by others. By confirming the current views or beliefs of those to whom they are communicating, the poster can enhance the credibility of the initial post. These tendencies to share inaccuracies beg the question of what personality traits drive the sharing of information that may not be factual?

Traits that Drive Misleading Online Information Sharing

As consumers interact in an online environment, some of them may be inclined to engage in malicious behavior. The overall propensities of an individual can profoundly influence their inclination toward such activities. For example, Buckels et al. (2014) discuss varying propensities that manifest in consumers online, such as trolling, acting destructively, or disruptively online setting with no apparent purpose. This type of malicious behavior is distinct from other online activities (such as chatting, debating topics, or Internet browsing) in that trolling is highly linked to the Dark Triad traits (Machiavellianism, Narcissism, and Psychopathy) (Buckels et al., 2014; Larson and Denton, 2014).

Machiavellianism

Machiavellianism, the tenacious pursuit of one's own objectives (Sparks, 1994; Calhoon, 1969), is linked to propensities toward manipulation, cunning, tactical planning, and cynicism (Jonason and Webster, 2010; Jones and Paulhus, 2014; Christie and Geiss, 1970; Hare and Neumann, 2008). The concept of Machiavellianism evolves from Niccolo Machiavelli, whose political strategies focused on "the ends justifying the means." Prior studies have linked the presence of Machiavellian traits to misbehavior intentions (Daunt and Harris, 2011). Additionally, Machiavellianism is highly related to the opportunistic behavior of consumers (Ro and Wong, 2012). For instance, within the online forums focusing on trolling, numerous respondents (40%) discussed the ability to influence others for one's agenda or beliefs purposely.

"They want credibility and respect and authority they have never earned"

"I know for me personally it's more about me standing up for what I think is right or what I believe in when I see someone post something I disagree with. It's to make me feel like I didn't just give sometime [sic] a pass when I seriously think or know someone is wrong."

Machiavellianism was the most common trait present in the observed online trolling behaviors. When posters feel that another person's agenda was obvious, they tend to scrutinize the original poster. The real power of posters with Machiavellian tendencies seems to be the ability to fly under the radar until the person is ready to reveal his or her intentions.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
6	<p>“In their essence Trolling just digs for emotion. Any emotion can be target really as you just manipulate. Often its [sic] just for a rise of anger. Because anger is the easiest to achieve. Or its [sic] just trying to get any kind of response not [sic] matter the emotion.”</p>	I-words (I, Me, My)	0	5.51
		Social Words	4.3	9.71
		Positive Emotions	2.2	4.57
		Negative Emotions	4.3	2.1
		Cognitive Processes	23.9	10.77
		Summary Variables		
		Analytic	64.7	55.92
		Clout	41.4	55.45
		Authenticity	9.3	55.66
		Emotional Tone	4.1	63.35

Figure 2.6 Poster 6 Summary

Poster 6 viewed the overtly negative nature of trolling to be quite complex and utilized cognitive processes (23.9% of total word usage), and analytical thinking (64th percentile) in communication with others to better understand it. Additionally, these posters applied a substantial amount of cognitive resources to understand the purpose and motives of others, as described by Poster 7.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
7	<p>“By definition the act of creating fake reviews infers directly to dishonesty which, again by definition, is immoral and unethical. The two parties are not at equal footing as one party is being deceitful by attempting to manipulate the other party to buy a product that isn't necessarily endorsed.”</p>	I-words (I, Me, My)	0	5.51
		Social Words	6.1	9.71
		Positive Emotions	8.2	4.57
		Negative Emotions	4.1	2.1
		Cognitive Processes	22.4	10.77
		Summary Variables		
		Analytic	92.8	55.92
		Clout	34.1	55.45
		Authenticity	33	55.66
		Emotional Tone	91.8	63.35

Figure 2.7 Poster 7 Summary

Respondents to the “Ask me Anything” revolving around a person who is incentivized to provide fake reviews on behalf of companies echo the perceived manipulation by those who engage in purposely fraudulent reviews. When posters are aware of the financial motives for behavior, then they are very quick to employ cognitive resources to argue or disprove the validity of the poster’s motives, as seen in Poster 7’s message.

These sentiments are also present in a subreddit forum, with 59 participants discussing the issue of fake reviews on Amazon. The original poster, Poster 8, communicated concerns about the presence of fraudulent reviews and provided a narrative to justify the discontent. The poster contributed stories and examples to reinforce the points that are made and strives to influence the other participants in the subreddit forum.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
8	“The sellers ask you to buy and then they refund (sometimes with extra, so a payment) via paypal. I used to review back when it was allowed and stopped when Amazon drove it underground, because I refuse to write a fake or paid review and will only review with my actual thoughts. Sadly, I'm in the minority, most 'coupon reviewers' took Paypal as soon as 'incentivised [sic] reviewing' was banned. Back in the days when I reviewed, you had to put a disclaimer, so at least customers were aware that it was a review in exchange for a product, and could choose to ignore it. Right now, the problem is consumers have no way of knowing which are fake.”	I-words (I, Me, My)	4.2	5.51
		Social Words	4.2	9.71
		Positive Emotions	0	4.57
		Negative Emotions	4.2	2.1
		Cognitive Processes	11	10.77
		Summary Variables		
		Analytic	48	55.92
		Clout	50	55.45
		Authenticity	67.5	55.66
		Emotional Tone	1	63.35

Figure 2.8 Poster 8 Summary

Here, the original poster seems to exhibit Machiavellian propensities as the purpose of the post, even though not necessarily for malicious intent, is to purposely influence the others to take his/her side on a topic. The poster attempts to convey high levels of authenticity to influence the

forum posters (68th percentile). Based on the comments posted on the forums, Machiavellian messages tend to focus on the creation of perceived authenticity of the content to influence others. Another point of note when discussing Machiavellianism is that the trait itself might not always be easily detectable since the propensity relies on cunning and tactical manipulation of others. Therefore, an individual can act to persuade others and may even be judged as having benign intent, although they may be operating with malicious intent.

Narcissism

Narcissism is the creation of a grandiose and superior perception of oneself (Jonason and Webster, 2010; Jones and Paulhus, 2014). These individuals exhibiting levels of narcissism may create a cloud of confidence to mask insecurity. A narcissist's perception of superiority to others can act as a facilitator to justify negative actions. When individuals post and engage in negative online information, they tend to paint themselves as the "protagonist." This action is not uncommon since word-of-mouth generation tends to be positive (i.e., the poster is painted in a positive light) and the transmission of the message tends to focus on painting the other party in a negative light (DeAngelis et al., 2012). Narcissism emerged as the second most observed trait (35%) since the posters tend to share their expertise and confidently share information to justify their positions. For instance, posters engaging in online trolling behavior convey that they are not the same as the others and convey themselves as a source of behavioral knowledge.

"I think it must be certain types of people that that happens to, because it doesn't happen to me, that's for sure."

I am a part time troll, for decades now. I'll let you into why I think it happens, from knowledge gathered in many places. Also, what you can do to not make yourself the target you are currently being.

Posters also shared stories that link them directly to an event and provide first-hand knowledge. This “self” focus tended to stem from narratives based around another party as the “antagonist” that the poster must address. For instance, posters voice how they have been wronged and how the other party is at fault. Posters may then provide “evidence” of expertise or experience to justify the information. Poster 9 below focuses on bringing attention to his/her own experience or connection to an event. The poster utilizes messages that attempt to convey clout, (87th percentile), to build credibility.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
9	“A family member of mine works for []. According to her employees are regularly reminded not to discuss what they do for work in public because [] is well aware of their terrible reputation and have, in response, basically abandoned working on customer service all together [sic].”	I-words (I, Me, My)	2.3	5.51
		Social Words	18.2	9.71
		Positive Emotions	2.3	4.57
		Negative Emotions	4.5	2.1
		Cognitive Processes	11.4	10.77
		Summary Variables		
		Analytic	74.3	55.92
		Clout	87.2	55.45
		Authenticity	19.8	55.66
		Emotional Tone	3.7	63.35

Figure 2.9 Poster 9 Summary

Psychopathy

Psychopathy is the presence of callousness and impulsivity in actions and is marked by a lack of remorse for one’s actions (Hare and Neumann, 2008; Jonason and Webster, 2010). Moreover, individuals with psychopathic tendencies are self-focused and have little regard for others. Many of the forums discussed this propensity openly. This impulsivity may also lead to emotional outbursts in information sharing and actions that may cause some level of harm to others but provides personal enjoyment at another’s expense.

“Uses it as an excuse to say awful things and then be like IF I SAY OUTRAGEOUS THINGS JUST ASSUME I'M JOKING. And then getting mad if I don't assume he's joking because he's made some awful joke.”

I then walked back to where I left it, noticed something was weird about it, popped it open, saw the papers, and had to COMPLETELY take it apart to fix it.

Underlying callousness toward others can justify negative online behavior. Although it is not as openly demonstrated as the other traits, the presence of Psychopathy provides insights into the justification by the individual in participating in the malicious behavior that appears to show enjoyment in the actions of harming other parties (25% of those engaging in malicious behavior).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
10	“Yup. Trolling is pretty much the use of a non-social environment (the Internet) to satisfy anti-social impulses without damaging one's social reputation. Those anti-social impulses could be caused by a number of things, but I think depression is a big factor.”	I-words (I, Me, My)	2.3	5.51
		Social Words	9.1	9.71
		Positive Emotions	4.5	4.57
		Negative Emotions	4.5	2.1
		Cognitive Processes	18.2	10.77
		Summary Variables		
		Analytic	94.8	55.92
		Clout	50	55.45
		Authenticity	5.4	55.66
		Emotional Tone	25.8	63.35

Figure 2.10 Poster 10 Summary

Based on the perceptions of this poster’s analytical breakdown of the behavior, psychopathic behavior (95th percentile) is described as more emotion-and impulse-driven, the response is discussed in a negative light among the posters since the overall emotional tone tends to be fairly negative (26th percentile). Other posters confirmed the presence of those posters that attempt to influence others by posting overtly negative tone and impulsive messages.

*“Still others think they have a chance to change others minds with a few misspelled words and/or *exclamations!!!”*

Posters who showed signs of psychopathic tendencies were overtly vocal in discussing what they liked but also what the disliked. These posters conveyed anger and sometimes violent tendencies when talking about hated brands or companies. They also exhibited high levels of disdain for companies that they viewed in a negative light by posting dark jokes and addressing the perceived absurdity of some company products.

“Right?! Like is this a joke? They have yet to respond to my angry Twitter messages”

There's no downside to watching it. You get to see an uncommon car and an explosion!

Additionally, posters seemed to have little inhibitions while addressing the brands and are quite open to sharing their viewpoints. For instance, Poster 11 below uses humor to convey hatred toward a company but also communicates authenticity (85th percentile) while sharing a negative emotional tone (26th percentile).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
11	<p><i>“Right??? OMG, don't even get me started on that dump. My hubs insists on going there every once in a while for some hardware this or that and by the time we get out of there I'm ready to punch a kitten in the face. Nothing on the shelves, dirty everything, mismarked prices, no price scanners and employees that literally look like they wish they were dead. Go away already!!!!”</i></p>	I-words (I, Me, My)	4.3	5.51
		Social Words	5.7	9.71
		Positive Emotions	1.4	4.57
		Negative Emotions	1.4	2.1
		Cognitive Processes	8.6	10.77
		Summary Variables		
		Analytic	74.2	55.92
		Clout	38.7	55.45
		Authenticity	84.7	55.66
Emotional Tone	25.8	63.35		

Figure 2.11 Figure 2.10 (continued)

Based on the observed discussions, online maliciousness seems to present itself when participants with greater propensities toward any of the Dark Triad see an opportunity and then utilize whatever resources available to establish credibility to meet their goal. Being able to rationalize the impulse for the negative actions themselves may be important in determining whether an individual engages in the behavior. Next, the emergent themes related to the motivation and traits that drive online misbehavior will be applied to an additional forum to provide evidence of prevailing motivations and traits involved in the practice. This application will determine whether they are present, supporting the validity of the findings.

Application of Themes to Forum Discussing Pricing Mistake

Reddit and many other online communities can attract avid followers who create forums in real-time to address a specific topic, issue, or opportunity. A Reddit forum, consisting of 229 participants, that focus on the purchase of technology for both the personal consumption and the resale of products was chosen to test for the themes uncovered in the previous forums. Once more, the online disinhibition effect appears to be a factor as the participants of the forum seem to work regular jobs and only collaborate based on pricing opportunities.

“I ordered that about 10 minutes ago from their site. Got the order confirmation before it went down, let's hope it ships!!”

The community frequently coordinates to educate buyers regarding price comparisons, special purchasing, and purchasing opportunities. The topics of the posts vary based on real-time access to new pricing information that is found through online searches. Specifically, the subreddit forum that served as the sample for comparison (1,900 posts, 229 participants) discusses a pricing mistake and the community's exploitation of the mistake afterward. The occurrence in the forum is distinct from a service failure, offerings that fall below a customer's expectations

(Hess, Ganesan, and Klein, 2003), in that many consumers did not expect to get pricing initially. The participants quickly took advantage of an obvious pricing error to purchase the products and then coordinated by sharing links to corresponding websites where the products were available.

"Bought 6 to use as a wall. Wife didn't approve... " casually asks for triple the price paid."

27" went through for me, fingers crossed for no cancellations.

"My order is still pending so I filed a complaint with Amazon and they are looking into it."

This process took place over a few days before the companies were made aware of the errors.

Proactivity in Pursuing Opportunities

The purpose of the subreddit forum under observation is to "link users to products that are on sale at various websites." Therefore, the intent for each user is to get "deal" pricing. However, the community quickly shared information regarding what many deemed to be a pricing error.

"Somebody probably lost their job at [company] today."

"You handle a price mistake that terribly, you probably shouldn't be owning your own business anyway. Survival of the fittest."

The Dark Triad traits were also evident in forum postings. Of those posters in the forum who exhibited the Dark Triad traits, Machiavellianism (47%) of the posts was apparent, tactical communication, and planning of purchases.

“I'm sure plenty of people despise what I do also, but damn if it isn't nice to supplement my day job earnings by grabbing deals that sometimes net like a 500%+ turnaround on the initial investment.”

“Yeah no matter how many times you contact them they just ask their supervisor who says there's nothing they can do. Zero accountability. They lie to get you to go away.”

Narcissism, (35% of the posts), manifested in the sense of self-importance and knowledge of company policies.

“I generally only browse and buy what I know. So I mainly troll deal sites and subreddits that align with my interests.”

Psychopathy (18% of the posts) appeared as posters exhibited sheer enjoyment in punishing the company that made the mistake and who was unable to honor the pricing.

“If they want to lie and swindle to protect their store, they shouldn't get away with it.”

The community members quickly shared information about the purchasing process and the response from numerous resellers. Many of these posters were hopeful that the purchase would be honored at the erroneous price, regardless of the error.

“Same [sic] mine changed to preparing for shipment, and also my credit card got charged , so keeping my fingers crossed.”

“The Big Lie here is "they cannot honor the price." They can, they are choosing not to.”

Therefore, the majority of the posters (45%) were focused on the financial opportunity of the error. While posters also demonstrated amusement (16%), entertainment (12%), power-seeking (12%), and mischievous enjoyment (16%) motivations. LIWC analysis supports the researcher's evaluation of the posters' motivations, as evidenced in the output for Poster 12 below.

Informational content conveys an analytical evaluation of the order process (89th percentile) and a highly positive emotional tone (99th percentile) that s/he will get what s/he wants.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
12	"I ordered the 27" one a couple hours ago. Prepping for shipping, not cancelled [sic] yet, and I already got charged for it...Hoping for the best...fingers crossed..."	I-words (I, Me, My)	7.1	5.51
		Social Words	0	9.71
		Positive Emotions	7.1	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	7.1	10.77
		Summary Variables		
		Analytic	88.7	55.92
		Clout	7.7	55.45
		Authenticity	52.9	55.66
		Emotional Tone	99	63.35

Figure 2.12 Poster 12 Summary

Poster 13 utilizes more clout-based messages to others by discussing his experience and the “luck in the process (65th percentile) and also conveys a highly positive tone (91st percentile).

The posters clearly seem to take enjoyment in exploiting the opportunity before them.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
13	"Didn't receive the email that everyone was getting from [] regarding backorders and that they'll have to wait. No messages from them on [] as well. I'm assuming that I'm one of the lucky ones that manage to nab one of the monitors while it was still in stock. Only time will tell..."	I-words (I, Me, My)	3.9	5.51
		Social Words	13.7	9.71
		Positive Emotions	3.9	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	7.8	10.77
		Summary Variables		
		Analytic	36.1	55.92
		Clout	65.3	55.45
		Authenticity	53.8	55.66
		Emotional tone	90.5	63.35

Figure 2.13 Poster 13 Summary

Although some posters seemed happy that a great deal might be achieved, others were vocal about their intentions to exploit the companies. For instance, some users aggressively discussed the opportunity presented by the pricing error and conveyed steps to engage in “predatory shopping.”

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
14	“I was way more into "predatory" shopping 10 years ago when it supplemented an income that was significantly lower than it is today, but old habits die hard I guess, and it's still the easiest money I know how to make.”	I-words (I, Me, My)	7.3	5.51
		Social Words	0	9.71
		Positive Emotions	2.4	4.57
		Negative Emotions	2.4	2.1
		Cognitive Processes	14.6	10.77
		Summary Variables		
		Analytic	21.2	55.92
		Clout	11.1	55.45
		Authenticity	99	55.66
		Emotional Tone	25.8	63.35

Figure 2.14 Poster 14 Summary

Poster 14 states that this practice is utilized to purposely find and exploit pricing errors that can then be used to reap a profit. The poster also provides steps for others to use to engage in the predatory shopping themselves. The LIWC output for this post heavily focused on building perceived authenticity in the message (99th percentile). Moreover, the poster appears to convey himself as a proponent of the other participants and begins to utilize the forum as a platform to voice vigilantism. Poster 15 stated the following in agreement. The poster suggested that all opportunities should be pursued if possible. The LIWC output shows that this post consisted of high levels of analytical thinking (96th percentile) and authenticity (79th percentile).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
15	“I get orders cancelled [sic] most of the time, but you miss out on 100% of the price mistakes you don't place an order for. I most recently bought a bunch of Blu-ray boxsets at 80-90% off retail. Thought for sure they'd all be cancelled [sic], but every item that was in stock shipped out.”	I-words (I, Me, My)	3.7	5.51
		Social Words	5.6	9.71
		Positive Emotions	1.9	4.57
		Negative Emotions	3.7	2.1
		Cognitive Processes	16.7	10.77
		Summary Variables		
		Analytic	95.8	55.92
		Clout	50	55.45
		Authenticity	79.4	55.66
		Emotional Tone	5.8	63.35

Figure 2.15 Poster 15 Summary

Vigilantism

The vigilantism theme that emerged from the previous forums becomes increasingly apparent again as posters voice levels of expertise beyond the typical boundaries of online purchasing.

“I'm still confused by "we respectfully request that you cancel your order...". What if I don't? It sounds like they don't want to cancel it themselves, right?”

“If this is FBA then they want you to cancel. If they cancel the order it affects their metrics.”

This vigilantism is then exacerbated by the diminished inhibition in the online environment.

First, the majority of participants (64%) quickly voiced concern and discontent,

“Oh I know. I wish they would honor it like they told me they would if I waited. I had no problem waiting 4-6 weeks to get it...But to lie to me...nope. Will not cancel. I will let them take the hit on Amazon and wait out the 30 days.”

discussed experiences,

“They said they generally get items shipped out very quickly and mine was already prepping for shipping within half a [sic] hour of the order.”

coordinated responses,

DO NOT cancel your order. Let them do that and take the reputation hit, they deserve that for stringing people along and lying,

and share email responses from the companies. Additionally, since the pricing mistake affects numerous companies, the participants compare company responses.

“Is there anyway [sic] people who didn't order from them could rep them? They deserve the positive ratings.”

“Yeah, why do they keep going with this statement? They're honest to one guy and not the next?”

The other 34% of posters communicated with the posters or posted tangentially in the discussion. However, vigilantism presented itself quite heavily as participants began sharing information that aimed to discredit the companies involved through explaining the review process, researching company information

“Yes, but the sale is long over, the item has been pulled at that price from amazon and the company adjusted their website.”

and voicing expertise associated with online pricing availability.

“And Reddit successfully crashed their website, the entire thing is down now lol.”

These posters were quite active in sharing their views and beliefs. The LIWC output for Poster 16 indicated that the poster used social words (10.9%), negative emotions (6.5%), analytical thinking (84th percentile), clout, (86th percentile), and a negative tone.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
16	"1 month later, [] finally admitted to a price mistake and cancelled [sic] after stringing us along with these fake excuses. Was it that hard in the first place? Made sure to give them a poor review (2 stars so it wouldn't be deleted like the 1 stars)."	I-words (I, Me, My)	0	5.51
		Social Words	10.9	9.71
		Positive Emotions	2.2	4.57
		Negative Emotions	6.5	2.1
		Cognitive Processes	10.9	10.77
		Summary Variables		
		Analytic	84.1	55.92
		Clout	86.2	55.45
		Authenticity	9.3	55.66
		Emotional Tone	1	63.35

Figure 2.16 Poster 16 Summary

Next, the posters voiced discontent associated with perceived actions by the companies after the fact by checking for shipping statuses,

"Pending transaction with my bank still hasn't gone through, and the order status is still sitting on "Preparing for Shipment". I think they're just going to let it time out on its own at this point."

checking to see if negative reviews have been removed,

"They removed it again but I got a screenshot of it"

and researching company information.

"It is not going to take down a giant billion dollar company like [company]"

The posters voiced the importance of being active to ensure that proper reviews are posted and that participants pursue the pricing and refuse to cancel their orders.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
17	"Got that email this morning...So first they send me an email that says to wait 4-6 weeks and they will send me my order and now they are saying they will not send me my order and telling me to cancel my own order....uhhhhhhhh yeah ok so you lie to my face then want me to help you out by canceling my own order? Nah."	I-words (I, Me, My)	14.7	5.51
		Social Words	22.1	9.71
		Positive Emotions	1.5	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	2.9	10.77
		Summary Variables		
		Analytic	10.3	55.92
		Clout	72.2	55.45
		Authenticity	99	55.66
		Emotional Tone	53.4	63.35

Figure 2.17 Poster 17 Summary

Poster 17 used a tongue in cheek approach when discussing the topic. This approach indicates that the forum community context can be vital in determining the meaning of information that is shared. Once again, the poster conveyed high levels of I-words and (14.7%) social words (22%). Also, the poster aimed to build support for the message through high levels of clout (72nd percentile) and authenticity (99th percentile). An important point to consider in unpacking this comment is that even LIWC may have difficulty in detecting levels of sarcasm and tone in the post. Note that the emotional tone is indicated as fairly neutral (53.4%) even though the poster is clearly being sarcastic.

Justification

Vigilantism enhances the likelihood of participants retaliating when the company is painted in a negative light. For instance, the majority of the respondents actively voice concern and discontent while claiming that the companies are inherently greedy or unethical. Some posters were quick to claim that the companies deserved the treatment. These responses tend to reinforce other negative reviews and influence other participants in the discussion.

“Looks like [company] is having negative reviews removed from their Amazon store.

There were five negative reviews earlier today and now only one...”

Participants do not seem to feel the guilt associated with their actions that could harm since the actions are deemed as justified based on the perceptions of the companies. When the customers began to reinforce and confirm biases through justification, they explained this justification to others.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
18	“They can fulfill it, its a multi billion [sic] dollar company. Face what you guys did and honor the orders!”	I-words (I, Me, My)	0	5.51
		Social Words	21.1	9.71
		Positive Emotions	5.3	4.57
		Negative Emotions	0	2.1
		Cognitive Processes	0	10.77
		Summary Variables		
		Analytic	21.4	55.92
		Clout	99	55.45
		Authenticity	1	55.66
		Emotional Tone	97.6	63.35

Figure 2.18 Poster 18 Summary

The LIWC output for the Poster 18 above indicates the presence of a justified negative response through the use of clout (99th percentile), an overwhelmingly positive emotional tone (98th percentile), and analytical justification.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
19	"It would absolutely make my day, if what happened to [] happened to these guys. Get your store removed because of too many negative reviews. (Normally I wouldn't suggest a negative review for a price mistake, but these guys deserve special treatment for their dishonest and in some cases lack of communication)."	I-words (I, Me, My)	3.9	5.51
		Social Words	11.8	9.71
		Positive Emotions	2	4.57
		Negative Emotions	2	2.1
		Cognitive Processes	21.6	10.77
		Summary Variables		
		Analytic	52.5	55.92
		Clout	50	55.45
		Authenticity	1	55.66
		Emotional Tone	25.8	63.35

Figure 2.19 Poster 19 Summary

Poster 19 also justified the response and did so by voicing a negative emotional tone directed at the company (26th percentile). Other posters also shared this distrust of the companies and negative perception.

"I doubt anyone is going to read this, but just a psa. Amazon is automatically deleting all one star reviews under the assumption they are fake. If you want to leave a negative review, try to leave 2 or preferably 3 stars."

Has anyone else noticed the downright defensive/hostile responses [company representative] on the 1 star reviews? In one she even mentioned that it totally wasn't a pricing glitch. My order is still pending so I filed a complaint with [company] and they are looking into it.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
20	“So, I just submitted a complaint to customer service about their stonewalling. feedback “So, I just submitted a complaint to customer service about their stonewalling, feedback manipulation, and basic dishonesty. Please take screen shots [sic] of their feedback page at various times in order to have proof of 1 star reviews being removed. I’m doing the same on my end.”	I-words (I, Me, My)	6.5	5.51
		Social Words	6.5	9.71
		Positive Emotions	2.2	4.57
		Negative Emotions	2.2	2.1
		Cognitive Processes	4.3	10.77
		Summary Variables		
		Analytic	80	55.92
		Clout	50	55.45
		Authenticity	54.9	55.66
		Emotional Tone	25.8	63.35

Figure 2.20 Poster 20 Summary

Additionally, like Poster 20, some posters were quite analytical (80th percentile) in assessing the situation and developing a response. The poster diligently shared information, responses, and actively coordinated with others to post 1-star reviews online.

Although the majority of the discussion in this forum took place in a negative light, not all participants promoted malicious retaliation. Some posters felt some level of guilt for the actions of the group.

“Is it weird that I don't want to order because I feel bad for [company]? This isn't Amazon, this is some vendor who probably lives and dies by the margins.”

However, Reddit contains a feature that allows users to sort responses. Responses that defended the companies were downvoted and were labeled as “controversial” meaning a user would have to scroll further down to see them at all. A poster in an earlier forum discussed this concern by stating the potential issues associated with only providing confirmatory information.

“Reddit is a hyper opinionated place. This makes for a lot emotional discussions. I think people enjoy seeing opinions they disagree with shut down harshly. It's very satisfying. Thus, those type of comments get upvoted. I think this harshness when it comes to dealing with unpopular opinions spawns a negative feedback loop when those on the "losing side" respond equally as harshly and so and so on. Creating this toxic environment in some areas of reddit”

Therefore, the availability of negative information far outweighed the counteracting positive information. This purposeful effective removal of contra information indicates that the discussions about the misleading nature of the brands could have been, in themselves, purposely misleading. Because of this occurrence, the vigilantism effect of posters was potentially enhanced by the justification provided in the groupthink dynamic. The LIWC output for posters 21 and 22 indicated similar summary output variable scores for authenticity (94th percentile) and clout (96th percentile).

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
21	“Love how some people jump on an obvious ERROR and then get upset when its [sic] not honored...Shameful. Some scoundrels up in here.”	I-words (I, Me, My)	0	5.51
		Social Words	8.7	9.71
		Positive Emotions	8.7	4.57
		Negative Emotions	8.7	2.1
		Cognitive Processes	21.7	10.77
		Summary Variables		
		Analytic	46.5	55.92
		Clout	50	55.45
		Authenticity	94.2	55.66
		Emotional Tone	25.8	63.35

Figure 2.21 Poster 21 Summary

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
22	"Downvote me all you want, but you guys are vultures. You are acting so entitled to this monitor. I love this sub and finding good deals but this was clearly an accident and you guys are being so cruel."	I-words (I, Me, My)	5.1	5.51
		Social Words	17.9	9.71
		Positive Emotions	5.1	4.57
		Negative Emotions	2.6	2.1
		Cognitive Processes	15.4	10.77
		Summary Variables		
		Analytic	1	55.92
		Clout	96.3	55.45
		Authenticity	3.7	55.66
		Emotional Tone	73.6	63.35

Figure 2.22 Poster 22 Summary

Poster 22 is proud to stand up using clout (96th percentile) to those who are sharing misleading information even though the dynamic supports the silencing of dissenting posters through downvoting. Although Poster 22 is deviating from the rest of the group, s/he does so with a positive emotional tone (74th percentile) while Poster 23 voicing negative emotions and a largely negative overall tone.

Poster Number	Quote	Traditional LIWC Dimension	Your data	Average for Social media: blog
23	"The entitlement these people have, this much negative feedback could ruin a seller."	I-words (I, Me, My)	0	5.51
		Social Words	7.7	9.71
		Positive Emotions	0	4.57
		Negative Emotions	7.7	2.1
		Cognitive Processes	7.7	10.77
		Summary Variables		
		Analytic	42.7	55.92
		Clout	50	55.45
		Authenticity	1	55.66
		Emotional Tone	1	63.35

Figure 2.23 Poster 23 Summary

This suggests that posters can actively attempt to mitigate the effects of information sharing and feel justified in doing so. However, since these posters were downvoted and labeled as “controversial,” posters new to the conversation would likely not see them. When the counterarguments were effectively hidden by downvoting, the availability of negative information far outweighed the positive. Thus, the posters tended to develop an echo chamber to reinforce beliefs, intent, and motive in pursuing the price. This ability to downvote information, can in itself display levels of Machiavellianism since the information stream is altered to fit an individual’s narrative.

A notable finding that was uncovered in the netnographic content analysis is the fact that the forum group has created a “Hall of Fame” page for the best deals ever retrieved. One of the nominations was the pricing error discussed by the posters above. Importantly, the group labeled the nomination as a “pricing mistake,” confirming that although individuals conveyed that the “unethical” companies created the animosity, participants were clearly cognizant that it was an actual pricing mistake and not just “a really good deal.”

The purpose of the netnographic content analysis was to uncover the motivations and traits of those who engage in malicious consumer behavior online. Several distinct themes (the Dark Triad, vigilantism, proactivity, and justification) emerged and were then supported by LIWC analysis. The findings of the netnographic content analysis provide essential insights into the traits of individuals who create and share misleading or malicious information with others. The participants in the communities reflected on not only the behavior of others but also what may be driving their behavior. Individuals in the forums analyzed in the follow-up analysis confirmed those insights.

Additionally, the netnographic content analysis, along with the *a priori* literature review, indicated that individuals who are motivated to seek out opportunities tend to also exhibit tendencies toward Machiavellianism, Narcissism, and Psychopathy. Even though present in most people to some extent, higher levels of these propensities tend to drive the pursuit of actions that benefit oneself at the potential expense of others (Jones and Paulhus, 2014). Posters then reaped the benefits of an online sharing platform since vigilante posters can spread information and take on leadership roles to coordinate others. Finally, the posters who create and share information with others appear to rationalize or justify to others, and themselves, what they are doing and why they are doing it. This finding provides important insights into not only the thought processes of a malicious poster but potentially the group rationale that can develop.

In conducting the qualitative study, a theoretical framework emerged and will be expanded upon in the following sections. The next section introduces the emerging theory from the netnographic content analysis and the conceptual development of the quantitative analysis. Study 2 will quantify the conceptual relationships that were uncovered through the netnographic content analysis by describing how a consumer may feel that sharing misleading information is justifiable. Driven by the Dark Triad, that consumer may work to ensure that their misleading posts have the maximum persuasive impact.

Inoculation Theory

Because the consumers involved in the online discussions are exposed to a large quantity of information of unknown quality, it follows that information manipulation could be at play (Dawson and Brashers, 1996; McCornack, Levine, Solowczuk, Torres and Campbell, 1992). Information can be manipulated through the level of quantity provided, the level of quality, the relation (or relevance to the topic at hand), and the manner (or presentation of the message)

(Dawson and Brashers, 1996). None of the research mentioned above explains the reasoning for the creation and dissemination of the fraudulent information in the first place. Inoculation theory provides this explanation since it describes not only how information is created but also how sharing a message with others can reduce the effectiveness of counterarguments from brands.

Inoculation theory, developed by McGuire (1968), uses concepts from vaccination and immunization literature to explain how individuals can become immune to certain information. In the physiological context, vaccines work by introducing a weak form of a virus so that the immune system can learn to fight it off. The theory applies the same logic to social communication and persuasion. For instance, the theory states that an individual, or entity, identifies some informational threat and then provides a weak preemptive argument (reputational preemption) (Banas and Rains, 2010). By identifying a potential threat early, the individual is then emboldened to protect a belief or attitude (Banas and Rains, 2010; Compton and Ivanov, 2012).

Companies have implemented inoculation policies to guide advertising campaigns (Szybillo and Heslin, 1973), such as Nestlé's response after scandals associated with baby formulas and food quality. Inoculation programs help to build resistance to external persuasion (Lessne and Didow, 1987) such as healthcare companies creating anti-smoking campaigns to discourage teenage smoking. Additionally, inoculation strategies can help to establish a resistance to service failures (Mikolon, Quaiser, and Wieseke, 2015) through the development of satisfaction and loyalty to offset negative experiences. Each approach listed above allowed a brand to prevent crises from occurring or acted to offset a current crisis. Although inoculation theory has typically been related to business communication and recovery strategies, individuals can act to self-inoculate or autoinoculate. However, how could this process take place?

Dual-process models describe how individuals can make decisions and have been applied in cultural development (Lizardo et al., 2016), trust development (Murray et al., 2011), moral judgments (Haidt, 2001), and prejudice (Dovidio, Kawakami, and Johnson, 1997). Inoculation may also develop through a dual-process approach. Posters who are influenced by individual proclivities, such as the Dark Triad, along with cognitive justification, begin to autoinoculate through this dual-process approach. Specifically, a poster can realize a belief (“I really want the pricing, and I deserve it”) and then justify the actions (the companies are liars, cheats, and they deserve to pay”) to create the counterargument through misleading information. Then when others are exposed to the said message that challenges current attitudes or beliefs, the receiver is more likely to disregard it (McGuire, 1968; Compton and Pfau, 2009).

In the context of the pricing error mistake, consumers perceive the response by the marketer as a threat to their personal goals of getting a discounted price and then autoinoculate through a confirmation bias that they deserve to get the incorrect price or that the company deserves negative treatment. The autoinoculation is then transmitted to the forum, group, or message board through the dissemination of a fraudulent or misleading message. Therefore, the initial bias or snap judgment against the firm then becomes viral, spreading to others. This two-process approach to inoculation suggests that the theory operates under the dual-process umbrella of decision-making.

Dual-process decision models indicate that decision-making is influenced by both automatic and analytical systems (Kahneman, 2011). For instance, a consumer may be driven by prior heuristics and even propensities that can impact the types of decisions that are made. These heuristics or snap judgments can be driven by personality traits, prior beliefs, or motivations. Consumers may also require the use of additional cognitive resources when determining which

actions to take. Therefore, the consumer may enact these additional resources to solidify a belief or justify a bias. When these individuals can consistently utilize both systems in judgment to create a congruent view of actions, they may be able to then build up immunity to contrary viewpoints.

Once the original poster has realized the threat (after autoinoculation) and provided refutation preemption (weak preemptive counterargument), an incubation period must be implemented (i.e., delay). If the original party offers too much information, others may become aware that the originator of the message is attempting to persuade them overtly, and the receiver of the message may then begin to question the validity of both statements (McGuire, 1964). Therefore, as the theory suggests, overtly persuasive statements can backfire, but too much of a delay in response can dilute the impact of the inoculation (Banas and Rains, 2010). Since online environments can act as a collective ecosystem, and other members can reinforce information sharing, the original poster does not necessarily need to enhance the original message. Prior studies have shown that inoculation can be long-lasting (Banas and Rains, 2010). The dynamic environment between posters can act to reinforce the incubation period required to strengthen corresponding views.

Finally, the overall involvement (personal relevance of the message) acts to embolden others to discredit the counterargument from the external party (Pfau, 1997). The original poster must communicate a message that is relevant to the receiver. Therefore, the original poster must utilize logical extensions to refute the counterargument while tapping into the beliefs, attitudes, and preferences of the receiver. When other posters get involved, the original poster may take on a less active role in sharing information while the new receivers of the message act to inoculate others. Since posters may have a shared “vested” interest in dissuading others from changing

attitudes, the group can develop a shared inoculation that reinforces current goals, much like herd immunity. This autoinoculation viewpoint provides a necessary comparison to the inoculation strategy implemented by companies. But how does the idea of autoinoculation hold up when applied to the themes developed in the netnographic content analysis?

The netnographic content analysis described that individual posters who have greater propensities and motivations (biases and beliefs) could be more likely to create and then share misleading, malicious information when justification (implementing cognitive resources) is present. In the next sections, insights from the qualitative analysis will inform the conceptual development of a model that will explain the antecedent traits that predict the dissemination of false information and the role that autoinoculation plays.

Conceptual Development of the Model

The Dark Triad

The Dark Triad encompasses an individual's propensity toward Machiavellianism, Narcissism, and Psychopathy and thus has been shown to influence online decision-making and behavior. Although the Dark Triad has been discussed as a "triad," most studies have measured the dimensions individually. For instance, Buckels et al. (2014) measure each component independently when looking at the propensities of online trolls. Jonason, Webster, and Schmitt (2009) analyze each dimension separately in addressing tendencies toward exploitative personalities. Additionally, Jonason and Webster (2010), with the "Dirty Dozen" scale, improved the measures for each piece of the Dark Triad (Machiavellianism, Narcissism, and Psychopathy) by creating a more parsimonious scale that measures each dimension. However, other researchers have summated the variables to create a total composite score (Rapp-Ricciardi, Widh, Barbieri, Amato, and Archer 2018; Jonason and Tost, 2010).

Conversely, this method may be deemed problematic since the overall composite score does not represent a reflective construct exhibiting unidimensionality (Lee and Cadogan, 2013). Although the Dark Triad constructs are statistically related, each construct has been measured for validity and reliability individually (Jonason and Webster, 2010; Jones and Paulhus, 2014). This section will provide both the theoretical justification to support the empirical specification for the Dark Triad as a first-order reflective and second-order formative construct (Coltman, Devinney, Midgley, and Venaik, 2008; Jarvis, MacKenzie, and Podsakoff, 2003). Therefore, it is crucial to establish the role of each construct, Machiavellianism, Narcissism, and Psychopathy, in forming an overarching formative Dark Triad construct. For that reason, the researcher will create a formative model construct to capture each reflective dimension.

Machiavellianism

Extant literature, establishing that Machiavellianism is linked to manipulation, cunning, tactical planning, and cynicism (Jonason and Webster, 2010; Jones and Paulhus, 2014; Christie and Geiss, 1970; Hare and Neumann, 2008), indicates that these traits will play a crucial part in forming the Dark Triad construct. Prior studies have linked the presence of Machiavellian traits to misbehavior intentions (Daunt and Harris, 2011), consumer ethics (Al-Khatib, Vitell, and Rawwas, 1997), cheating on service guarantees (Wirtz and Kum, 2004), and the ability to persuade others (Langner, Hennigs, and Wiedmann, 2013). Since *a priori* justification is present in establishing the role of Machiavellianism as a component of the Dark Triad, the propensity is expected to form a dimension of the formative Dark Triad construct.

Narcissism

Narcissism is the creation of a grandiose and superior self-perception (Jonason and Webster, 2010; Jones and Paulhus, 2014). Also, narcissists may create a cloud of confidence to mask insecurity. When a narcissist's perception of superiority is used to influence others, it can act to contribute to the justification of negative actions. For instance, consumers with narcissistic tendencies can feel entitled (Boyd and Helms, 2005), focus on self-image (Berthon, Pitt, and DesAutels, 2011), and develop consumption identities through narcissistic individualism (Cherrier, 2009). When individuals post and engage in negative online information, they tend to paint themselves as the "protagonist" (DeAngelis et al., 2012). Since *a priori* justification is present in establishing the role of Narcissism as a component of the Dark Triad, the propensity is expected to form a dimension of the formative Dark Triad construct.

Psychopathy

Psychopathy is the presence of callousness and impulsivity in actions and is marked by a lack of remorse for one's actions (Hare and Neumann, 2008). Moreover, individuals with psychopathic tendencies are self-focused and may engage in cruel practices (Karampournioti, Hennigs, and Wiedmann, 2018), brand bullying (Breitsohl, Roschk, and Feyertag, 2018), and consumer fraud (Harrison, Summers, and Mennecke, 2016). Since Psychopathy is typically marked by impulsiveness and emotional outbursts in information sharing, it is expected to influence callousness toward others. Since *a priori* justification is present in establishing the role of Psychopathy as a component of the Dark Triad, the propensity is expected to form a dimension of the formative Dark Triad construct.

Hypothesis 1a: Machiavellianism will significantly form a dimension of the Dark Triad construct.

Hypothesis 1b: Narcissism will significantly form a dimension of the Dark Triad construct.

Hypothesis 1c: Psychopathy will significantly form a dimension of the Dark Triad construct.

Proactivity in Pursuing Opportunities

Proactivity, a proclivity to pursue actions that influence one's current state, is a positive trait in both service providers (Bateman and Crant, 1993) and consumers (Kim and Rucker, 2012) in reducing harmful incidences. However, individual differences can drive the direction of this proactivity, whether it be toward benevolence, through improving one's personal or family wellbeing, or toward malevolence, taking advantage of others for one's own selfish gain. Consumers who exhibit propensities toward the Dark Triad have been shown in previous studies to exploit others (Jonason, Webster, and Schmitt, 2009). This exploitative nature can tend to involve a person's proactivity toward aggression (Jonason, Duineveld, and Middleton, 2015) and proactivity toward innovation (Wisse, Barelds, and Rietzschel, 2015). Proactivity toward innovation Wisse et al. (2015) can also influence the individual's propensities toward seeking out opportunities for success (Seibert, Kraimer, and Crant, 2001). When consumer proactivity is driven by negative affect, the corresponding actions can be quite harmful in seeking out resolutions to given problems (Bechwati and Morrin, 2003). Since negative forces have been shown to drive proactivity, it is expected that the Dark Triad formative construct will influence the consumer's proactivity in pursuing opportunities.

Hypothesis 2: The Dark Triad will positively influence a consumer's level of proactivity in pursuing opportunities.

Vigilantism

The concept of vigilantism, prompted by a need to transfer opinions to others (Saucier and Webster, 2010), can be used to describe actions that consumers can take to educate or inform others (Kravetz, 2007). For instance, consumers may act, as “watchdogs,” to ensure that the information posted in online reviews is accurate (Larson and Denton, 2014). Moreover, the pursuit of vigilance can act as an extension of consumer advocacy through the monitoring of pricing information and policies of brands (Wakefield and Inman, 1993). However, potential damage can be done when this vigilantism is prompted by a need to share opinions with others, without regard for accuracy (Saucier and Webster, 2010). This pursuit is marked by the view that the consumer’s opinion is the only accurate view, and it is their responsibility to “inform” others at all costs (Saucier et al., 2014).

Additionally, online resources allow consumers to act more proactively (Schivinski and Dabrowski, 2016) and serve as vigilantes, independent of marketers (Muniz and Schau, 2007). Because of this effect, proactivity in pursuing opportunities will positively influence a consumer’s level of vigilantism.

Hypothesis 3: The consumer’s level of proactivity in pursuing opportunities will positively influence the consumer’s level of vigilantism.

Sharing Self-Created Misleading Information

When vigilantism is coupled with technological outlets such as social media (i.e., Reddit, Facebook, Twitter), consumers who are inclined toward vigilantism have a powerful outlet for message sharing. Consumers can use vigilante tendencies to create messages that inflict harm (Zolfagharian and Yazdanparast, 2017) while justifying that the message is “for the greater good” (Saucier and Webster, 2010). These individuals are active in posting negative information

to induce a brand to “give in” to demands (Grégoire, Legoux, Tripp, Radanielina-Hita, Joireman, and Rotman, 2018). Additionally, consumers can then act to create information to influence others with a particular belief or preference. When this occurs, the individual uses vigilantism as a means to further communicate and persuade others in accepting a viewpoint (Saucier and Webster, 2010). Therefore:

Hypothesis 4: The consumer’s level of vigilantism will positively influence the likelihood of sharing self-created misleading information.

Dark Triad and Proactivity

Trolling behaviors, the practice of behaving in a destructive or disruptive manner in an online setting with no apparent purpose, rely on deceptive practices (Buckels et al. 2014; Binns, 2012; Coles and West, 2016). However, the deception itself is not always immediately apparent. For instance, an internet “troll” may present themselves as a jilted customer of a company in an attempt to sway online opinion or to create a crisis. This type of behavior is strongly linked to the Dark Triad traits (Machiavellianism, Narcissism, and Psychopathy) (Buckels et al., 2014; Larson and Denton, 2014). When individuals have a higher proclivity toward the Dark Triad, the level of proactivity can lead to innovative ways to solve problems (Wisse, Barelds, and Rietzschel, 2015). However, what means can a poster utilize when seeking out these opportunities? Proactivity in pursuing an opportunity can manifest itself when a consumer’s propensity toward vigilantism facilitates that opportunity. When communicating online, a consumer may voice complaints and spread negative word of mouth against brands, causing a firestorm (Dechêne, Stahl, Hansen, Wänke, 2010; Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). However, consumers may mask true intentions or motives while still acting as a match to spark a firestorm against a brand, regardless of message truthfulness. When a message is created, driven by the

Dark triad propensities, the proactivity in seeking out an opportunity can be based on malice, while being conveyed as advocacy for other consumers. The Dark Triad also drives a greater enjoyment in the misfortune of others, schadenfreude (Porter, Bhanwer, Woodworth, and Black, 2014). This personal gain at the expense of a brand and other consumers can be quickly vocalized to a larger group of consumers online. Since vigilantism may provide the drive to create and share information to implement this proactivity, it can be an important next step for individuals to voice ideas to and for others (Saucier and Webster, 2010; Grégoire et al., 2018). Vigilantism is expected to be a useful outlet for individuals with proactivity in pursuing opportunities.

Hypothesis 5: The consumer's level of proactivity in pursuing opportunities will mediate the relationship between the Dark Triad and vigilantism.

Proactivity and Vigilantism

Proactivity in pursuing opportunities can lead to an individual's likelihood of taking actions that benefit themselves. For instance, the proactivity toward aggressive or innovative behavior can be implemented when a consumer believes they have something to gain from another party (Jonason, Duineveld, and Middleton, 2015; Wisse, Barelds, and Rietzschel, 2015). This propensity, in and of itself, may be insufficient in addressing the overall likelihood of sharing misleading information. However, vigilantism provides the mechanism to implement ideas and influence other parties (Pfau, 2009) since its purpose revolves around propagating beliefs to others (Saucier, Webster, Hoffman, and Strain, 2014). Consumers who exhibit greater proactivity associated with pursuing desires, act more vigilantly in communicating with others. Since the presence of vigilantism can facilitate the consumer's intentions, the relationship

between proactivity in pursuing opportunities related to sharing self-created information may be dependent on vigilantism. This leads to the following hypothesis:

Hypothesis 6: The consumer's level of vigilantism will mediate the relationship between proactivity in pursuing opportunities and the likelihood of sharing self-created misleading information.

The Dark Triad, Proactivity, and Vigilantism

A consumer's propensity toward the Dark Triad construct is expected to influence the development of the inherent inclinations to create and share information and begins forming the required components of the automatic components of autoinoculation. An individual's automatic propensities provide a starting point for actions and simple decision-making (Kahneman, 2011). Therefore, a gut reaction, heuristics, or individual characteristics can drive the types of opportunities that are pursued since less cognitive resources are required to act (Evans and Stanovich, 2013). For example, if an online poster desires to mislead or harm others online because it is funny, the Dark triad tendencies (automatic propensities) then drive the level of proactivity in identifying this opportunity. Thus, vigilantism becomes import in sharing a viewpoint with others since it allows posters to do so in the face of potential conflict (Sonnentag, and Barnett, 2016).

Since the relationship between the Dark Triad and vigilantism is expected to be mediated by proactivity, and the relationship between proactivity and the likelihood to share misleading information is expected to be mediated by vigilantism, it is also predicted that proactivity and vigilantism will mediate the relationship between the Dark Triad and the likelihood of sharing misleading information with others. These relationships influence the more automatic components associated with the decision to act. Therefore:

Hypothesis 7: The consumer's proactivity in pursuing opportunities and the level of vigilantism will serially mediate the relationship between the Dark Triad and the likelihood to share self-created misleading information.

Justification (Autoinoculation)

Cognitive justification is a fundamental component of inoculation in that an individual must assess and justify the presence of an external informational threat. Thus, it is expected to influence an individual's ability to inoculate themselves. For instance, individuals justify their actions using accessible cognitive resources (Kim, Kim, and Park, 2012). When an individual justifies the feelings felt or motives realized, one can feel more comfortable with engaging in activities that may be deemed undesirable by others. Moreover, justification of actions helps individuals to maintain positive affect while bending the truth (Schweitzer and Hsee, 2002), pursuing purchases (Okada, 2005) accepting prices (Choi, Li, Rangan, Chatterjee, and Singh, 2014), self-persuading (Bernritter, van Ooijen, and Müller, 2017), and sharing online C2C information in illicit markets (O'Sullivan, 2015). Additionally, when a consumer deems that their actions are justifiable, by confirming an existing view or bias, those actions are more likely to be pursued (Malaviya and Sivakumar, 2002). The role of justification in enhancing the likelihood of sharing misleading information will require additional cognitive resources to allow for autoinoculation. Thus, justification is expected to strengthen the mediated relationship between proactivity in pursuing opportunities and the likelihood of sharing misleading information through vigilantism. Therefore:

Hypothesis 8: The level of justification (autoinoculation) will strengthen the mediated relationship between proactivity and the likelihood of sharing self-created misleading information through the consumer's level of vigilantism.

The mediated relationship between the Dark Triad and the likelihood of sharing misleading information focuses on more automatic decision-making and will likely be influenced by the more analytical (or logical) components implemented in justification (Bernritter, van Ooijen, and Müller, 2017). Thus, justification, though autoinoculation, is expected to strengthen the mediated relationship between the Dark Triad and the likelihood of sharing misleading information through proactivity in pursuing opportunities and vigilantism (all hypothesized relationships are available in Table 2.1). Therefore:

Hypothesis 9: The level of justification (auto-inoculation) will strengthen the serially mediated relationship between the Dark Triad and the likelihood to share self-created misleading information through the consumer’s proactivity in pursuing opportunities and the consumer’s level of vigilantism.

See the full conceptual model in Figure 2.24 below.

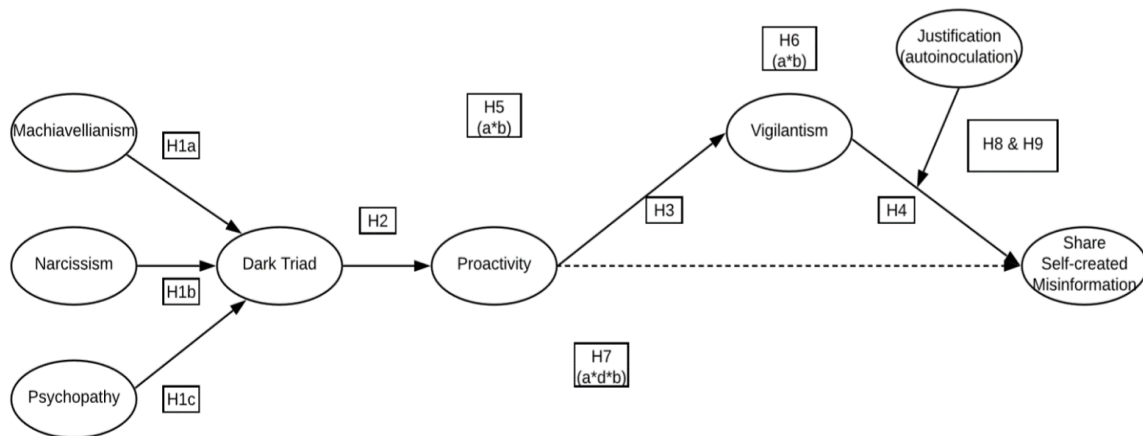


Figure 2.24 Conceptual Model

Table 2.1 Hypotheses

Formative Indicators	H1a: Machiavellianism=> Dark Triad
	H1b: Narcissism => Dark Triad
	H1c: Psychopathy => Dark Triad
Structural	H2: Dark Triad => Proactive
	H3: Proactive =>Vigilantism
	H4: Vigilantism =>Share Self-created Misleading Information
Mediation	H5: Dark Triad => Proactive => Vigilantism
	H6: Proactive =>Vigilantism =>Share Self-created Misleading Information
	H7: Dark Triad =>Proactive=>Vigilantism =>Share Self-created Misleading Information
Moderated Mediation	H8: Proactive * Vigilantism X Justify => Share Self-created Misleading Information
Moderated Mediation	H9: Dark Triad * Proactive * Vigilantism X Justify => Share Self-created Misleading Information

Quantitative Research

Study 2: Testing the Model

The constructs were measured using 7-point Likert-type and semantic differential scale items, which were adapted from previous research to 1) to capture the most parsimonious means to measure the constructs of interest and 2) to capture the unique dynamics of online information sharing. The scale items presented both reliability and validity in previous studies and are deemed appropriate for use in the current model. A pre-test was first conducted to determine the validity and reliability of the items as applied to the present study.

Measures Pretest

To ensure the reliability and validity of the adapted Likert-type scales and semantic differential scales, the researcher conducted a pre-test with a sample of 374 respondents on Amazon’s MTurk. Respondents with a 95% work acceptance rate were considered for participation, were compensated to reduce misrepresentation, and a working panel of

respondents approved from previous studies, and pre-tests were used (Hulland and Miller, 2018; Sharpe Wessling, Huber, and Netzer, 2017).

The total pre-test sample consisted of 50% female respondents, with the majority of respondents falling below 40 years of age (68.2%). Additionally, the respondents indicated that they often or exclusively purchase online (70.5%), use social media at least daily (79.4%), and participate actively in online discussions (54.3%). The respondents were compensated in exchange for participation in the study. A random subsample of 200 was taken to analyze the factor structures of the selected items. Acceptable reliability ($\alpha > 0.70$) (Netemeyer, Bearden, and Sharma, 2003; Peterson, 1994) was found for all measures. The pre-test results indicated the presence of some ambiguous items. Thus the scales were refined by using the top-performing items for both proactivity and vigilantism (Anderson and Gerbing, 1988; Hausman and Siekpe, 2009; Velicer and Fava, 1998). The CFA model was tested again on the remaining 174 respondents and then the total sample of 374 respondents. The measurement model maintained a good model fit in each. The items used in the measurement pre-test are tested again in the model pre-test to ensure that the fit is not sample-specific before items are used in the final structural model. Reliabilities and factor loadings for all items used in the final study to come are found in Table 2.2

Table 2.2 CFA Results, t-values, and Composite Reliability

Scales	c.r.	Standardized Factor Loadings	t-values
Machiavellianism – (Jonason and Webster, 2010)	0.93		
I tend to manipulate others to get my way.		0.90	25.32
I have used deceit or lied to get my way.		0.83	17.71
I have used flattery to get my way.		0.78	16.19
I tend to exploit others towards my own end.		0.98	*
Narcissism – (Jonason and Webster, 2010)	0.90		
I tend to want others to admire me.		0.76	15.33
I tend to want others to pay attention to me.		0.85	14.76
I tend to seek prestige or status.		0.85	*
I tend to expect special favors from others.		0.87	15.48
Psychopathy – (Jonason and Webster, 2010)	0.91		
I tend to lack remorse.		0.92	25.36
I tend to be unconcerned with the morality of my actions.		0.94	*
I tend to be callous and insensitive.		0.91	24.75
I tend to be cynical.		0.55	9.55
Proactive personality - (adapted from Bateman and Crant 1993)	0.84		
I enjoy facing and overcoming obstacles to my ideas.		0.75	12.25
I am great at turning problems into opportunities.		0.87	*
No matter the odds, if I believe in something I will make it happen.		0.66	10.57
I love being the champion of my ideas, even against others' opposition.		0.71	11.47
Justifiability of the decision - (adapted from Inman and Zeelenberg, 2002)	0.93		
Strongly Justifiable		0.90	23.41
Very Logical		0.75	15.65
Acceptable		0.92	24.93
Reasonable		0.93	*
Vigilantism (adapted from Saucier and Webster, 2010)	0.92		
I feel as if it is my duty to enlighten other people.		0.87	19.29
I feel that my ideas should be used to educate others.		0.89	20.19
I feel a social obligation to voice my opinion.		0.91	*
Sharing information (misbehavior intention) (adapted from Daunt and Harris, 2011)	0.95		
I would spread the information even though there is no genuine problem.		0.91	25.10
I would share exaggerated information when discussing the company.		0.90	24.01
I would share the unverified information to other customers.		0.93	*
I would share information in a way that the company may find unacceptable.		0.91	25.37
I would say overall, I can take advantage of others if it helps me meet my needs since people believe in me _b			
Know how to influence others regardless of the consequences because I am the center of attention _b			

Model fit statistics

Note: Chi-Square = 639.85, df = 298, p<0.01, Chi-square/df = 2.15, incremental fit index (IFI) = 0.95, Tucker-Lewis Index (TLI) = 0.94, comparative fit index (CFI) = 0.94, root mean square error of approximation (RMSEA) = 0.06.

All factor loadings have a p-value < 0.01

* denotes a constrained relationship to 1.00 for identification

b denotes reflective items used for formative model specification

Model Pretest

The model and hypotheses were next pretested with the final items rendered from the measures pre-test using the survey research method. In total, 502 respondents were recruited from a student sample pool using a snowball sampling technique. The snowball sampling technique has been implemented successfully to reach a larger and possibly a more representative sample of the population (c.f. Pagani, Hofacker, and Goldsmith, 2011; Ganesh, 1997). The respondents were given partial course credit for participating in the study and then sharing the survey link with three other respondents. The total model pre-test sample consisted of 53.8% male with a mean age of 32 years. Additionally, the respondents indicated that they often or exclusively purchase online (66.7%), use social media at least daily (75.2%), and participate actively in online discussions (58.3%).

The measurement model showed acceptable model fit based on the goodness of fit statistics (IFI, CFI, and TFI) which exceeded 0.95, composite reliability for each construct greater than 0.70, the AVE for each construct exceeding (.50) and exceeding the shared variance between constructs (Hu and Bentler, 1999; Garver and Mentzer, 1999; Fornell and Larcker, 1981). The results of the latent common factor analysis for common method bias indicated that a common method factor did not significantly change the model fit ($\Delta\chi^2/df1 < .01$).

The full structural results provided acceptable evidence of model fit based on the recommendations of Marsh, Hau, and Wen (2004) since the IFI, TLI, and CFI each exceeded 0.90. Next, the hypotheses from the formative, structural, mediation, and moderated mediation analyses were tested using a full structural model. The results for each hypothesis were significant and supported the contribution of the research. This pretest finding provides preliminary evidence of the presence of the relationships uncovered within the netnographic

content analysis. However, to ensure that the findings did not capitalize on the chance of using one dataset, an additional sample was retrieved and used to confirm the findings.

Sample and Procedure

The study implements a survey research methodology to test the hypothesized relationships. The researcher recruited 250 respondents for the confirmatory study. To ensure quality online respondents, the study included various attention checks, respondent screening policies, and a requirement of written text. Also, only U.S. respondents with a 95% work acceptance rate were considered for participation, a nominal wage was paid to reduce misrepresentation, and a working panel of respondents approved from previous studies, and pre-tests were used (Hulland and Miller, 2018; Sharpe Wessling, Huber, and Netzer, 2017). Six responses were dropped for failure to pass attention check measures through listwise deletion (Roth, 1994; Hair, Black, Babin, and Anderson 2010). The online consumer sample for the primary study included 244 respondents recruited from M-Turk who had purchased from companies online (Peterson and Merunka, 2014).

Participants were exposed to a scenario that provided a shopping situation in which a consumer has the opportunity to exploit a pricing mistake. The scenario captures the resulting justification in sharing the information with consumers online by stating that it would increase the likelihood that an order would be honored if others were complaining (see APPENDIX A). After exposure to the condition, respondents completed an online questionnaire (via Qualtrics). The sample consisted of 53% female respondents with a mean age of 39 years. 67.5% of the respondents indicated that they purchase either frequently or exclusively online, use social media at least daily (73.4%), and participate actively in online discussions (42.6%). (See Table 2.3).

Table 2.3 Sample Description

Age	Measures Pre-test	Model Pre-Test	Final Sample
18-20	2 (1%)	6 (1%)	2 (1%)
21-30	137 (37%)	246 (49%)	78 (32%)
31-40	116 (31%)	127 (25%)	77 (32%)
41-50	67 (18%)	68 (14%)	39 (16%)
51-60	33 (9%)	36 (7%)	26 (11%)
61 and over	19 (4%)	19 (4%)	22 (9%)
Gender			
Male	187 (50%)	270 (54%)	114 (47%)
Female	187 (50%)	232 (46%)	130 (53%)
Online Shopping			
Rarely	19 (5%)	31 (6%)	4 (1%)
Sometimes	91 (24%)	136 (27%)	56 (24%)
Often	212 (57%)	267 (53%)	137 (58%)
Always	52 (14%)	66 (13%)	40 (17%)
Social Media Usage			
Never	13 (4%)	13 (3%)	6 (2%)
Monthly	19 (5%)	39 (8%)	18 (7%)
Weekly	45 (12%)	70 (14%)	41 (17%)
Daily	251 (67%)	313 (62%)	143 (59%)
Hourly	46 (12%)	66 (13%)	36 (15%)
Social Media Discussion Activity			
Not very active	43 (12%)	44 (9%)	30 (12%)
Not active	78 (21%)	84 (17%)	58 (24%)
Neither active nor inactive	50 (13%)	81 (16%)	52 (21%)
Active	165 (44%)	237 (47%)	84 (34%)
Very active	38 (10%)	55 (11%)	20 (8%)

A manipulation check ensured that the participants were thinking of an online situation by asking them to “please think of how they would respond in an online setting,” “how the items describe them in an online setting” and specifically asking them to state “what occurred in the scenario and where explicitly.” The online consumer sample was compared to another sample of respondents that received an identical offline or “in real life” scenario rather than the online scenario. An independent samples t-test confirmed that the online purchase sample significantly

stated that the purchasing situation was taking place online ($t=25.86$ $p<0.01$). Therefore, the manipulation check was successful.

The Online Disinhibition Effect

Since individuals are expected to behave differently in an online versus offline setting, the impact of the online disinhibition effect is measured before reviewing the results of the online consumer group. This phenomenon is measured by comparing the variables of interest after the treatment is shown to respondents (the role of individual vigilantism on the likelihood to share self-created misleading information) in an online versus offline environment scenario. The χ^2 difference test for the online/offline groups was significant for the vigilantism-sharing information structural path ($\chi^2 = 5.74$, $p < .05$). Therefore, the two groups are significantly different. By comparing the standardized coefficients for the online group (.59) with the offline group (.34), it is evident that the online group has a significantly stronger standardized coefficient and is more likely to share the self-created misleading information.

The relationship between the Dark Triad and proactivity toward opportunities is non-significant in the offline consumer group when the direct path to vigilantism is introduced. Additionally, the indirect effect from the Dark Triad to the likelihood of sharing misleading information is non-significant for the offline consumer group. This finding indicates that those who have vigilante tendencies are more likely to share self-created misleading information in an online setting than in an offline or an “in real life” setting.

Confirmatory Factor Analysis

Tests using AMOS 24 showed that the composite reliability for all constructs exceeded 0.80, indicating evidence of construct reliability (Garver and Mentzer, 1999). The CFA fit

statistics for online consumers show a satisfactory model fit. The χ^2 of 639.85 with 298 degrees of freedom ($p < 0.01$) provides a χ^2/df ratio of 2.15. The IFI = 0.95, TLI = 0.94, CFI = 0.94 and RMSEA = 0.06, each providing additional evidence of acceptable model fit (Garver and Mentzer, 1999; Marsh, Hau, and Wen, 2004). Also, the t-values (found in Table 2.2) are all significant ($p < 0.01$). The AVEs for each construct exceed 0.50, providing evidence of convergent validity (See Table 2.4). The AVEs for each construct exceeded the level of shared variance between constructs, providing evidence of discriminant validity (Fornell and Larcker, 1981). The potential for common method variance was addressed by introducing a latent common method factor (Podsakoff, MacKenzie, Lee, and Podsakoff, 2003) to ensure that the measures are adequately represented. The results indicated that the presence of the latent common method factor did not significantly change the model fit ($\Delta\chi^2/df1 < .01$). Therefore, common method bias does not appear to be a major concern.

Table 2.4 Means, Standard Deviations, AVEs, and Shared Variance

	Mean	SD	1	2	3	4	5	6	7
Machiavellianism	3.06	1.74	0.76						
Narcissism	3.57	1.61	0.46	0.69					
Psychopathy	2.80	1.64	0.59	0.37	0.71				
Proactive	5.13	1.15	0.01	0.03	0.01	0.57			
Justify	3.71	1.86	0.24	0.17	0.27	0.01	0.77		
Vigilante	3.79	1.76	0.12	0.22	0.15	0.06	0.18	0.79	
Misinformation	3.05	1.80	0.38	0.32	0.45	0.03	0.59	0.30	0.83

Note: Average Variance Extracted is listed in the diagonals of the table

Structural Equation Modeling

Structural equation modeling provides flexibility in model creation, can reduce bias in measurement, and provides model fit statistics while calculating the moderated mediation effects when proper measurement and structural models are specified (Hayes, Montoya, and Rockwood,

2017). After the assessment of the measurement model, the full model was analyzed using AMOS 24 (see Table 5). The results of the structural model provide evidence of an acceptable model fit ($\chi^2 = 1023.96$, $df = 441$, $\chi^2/df = 2.32$, $p < 0.01$, IFI = 0.92, TLI = 0.91, CFI = 0.92 and RMSEA = 0.07) (Garver and Mentzer, 1999; Marsh et al., 2004; Schumaker and Lomax, 2004). Therefore, the structural model, mediation, and moderated mediation analyses can be confidently conducted.

The Dark Triad Formative Model

Since the reliability of a formative construct is calculated differently than that of reflective indicators, the recommendations of Diamantopoulos and Winklhofer (2001) are followed. The dimensions of the formative construct have been fully specified based on *a priori* expectations since numerous studies have discussed the specific implications of each component (c.f. Jonason and Webster, 2010; Jones and Paulhus, 2014). The number of indicators required to form the model is present in the current data by providing reflective items for specification in the formative construct (Diamantopoulos and Winklhofer, 2001). Multicollinearity is addressed by conducting a linear regression and isolating each independent (forming) construct individually (Hair, Black, Babin, and Anderson 2010). The VIF for all summated constructs was less than 4, indicating that multicollinearity was not an issue (Hair, Black, Babin, and Anderson 2010). Since the AVE for the construct is higher than the shared variance, discriminant validity is present, and the items should be kept (Fornell and Larcker, 1981). The steps recommended by Jarvis, MacKenzie, and Podsakoff (2003) for implementing the MIMIC formative model specification are used by creating two reflective items to capture the total Dark Triad construct (see Table 2). Additionally, this study provides an outlet for external validity to be established by utilizing prior literature to develop *a priori* expectations that are then extended to create the formative

construct. The formative relationships are analyzed with the remaining structural paths to better articulate the relationship between the measures, formative construct, and endogenous variables.

The measurement model is used to test the significance of Machiavellianism, Narcissism, and Psychopathy in forming the latent Dark Triad construct. The unstandardized coefficient for the path from Machiavellianism to the Dark Triad (0.45) is significant, $t = 4.78$, $p < 0.01$.

Therefore, Hypothesis 1a is supported. Machiavellianism is a significant component in forming the Dark Triad formative construct. The path from Narcissism to the Dark Triad is also significant (0.34, $t=4.54$, $p < .01$) and provides evidence of the role of Narcissism as a significant component in forming the Dark Triad. Therefore, Hypothesis 1b is supported. Finally, Psychopathy also has a positive and significant relationship in forming the latent Dark Triad formative construct (.18, $t=2.08$, $p < .05$) and supports Hypothesis 1c. These three components significantly form the whole Dark Triad construct. When consumers experience higher levels of Machiavellianism, Narcissism, and Psychopathy, greater overall propensities toward the total Dark Triad are realized.

The Dark Triad and Proactivity

The formative Dark Triad construct is used to assess the relationship with proactivity in pursuing opportunities. The unstandardized coefficient for the path from the Dark Triad (0.14) is significant, $t = 3.93$, $p < 0.01$. Therefore, Hypothesis 2 is supported. The Dark Triad significantly influences the level of consumer proactivity. A post hoc test was conducted to determine the potential influence of each of the Dark Triad components (Machiavellianism, Narcissism, and Psychopathy) individually on consumer proactivity in pursuing opportunities. Machiavellianism (-0.08, $t=-1.08$, $p=.28$), Narcissism (0.11, $t=1.72$, $p=.09$), and Psychopathy (.10, $t=1.42$, $p=.15$) were all non-significant in individually driving the development of proactivity in pursuing

opportunities. A second post hoc test compared the effect size of the formative Dark Triad construct with a composite Dark Triad variable. The Dark Triad composite variable was significant (0.03, $t=2.54$, $p<.05$). However, the standardized coefficient (.17) is weaker than the formative Dark Triad construct (.29). Therefore, the formative overarching measure explains more of the overall variance in the proactivity to pursue opportunities.

Proactivity and Vigilantism

The structural model was used to test the significance of the level of consumer proactivity in influencing vigilantism. The unstandardized coefficient for the path from proactivity to vigilantism (0.32) is significant, $t = 4.28$, $p<0.01$. Thus, Hypothesis 3 is supported. Consumers who are more proactive in seeking out opportunities are more likely to engage in vigilantism to influence others.

Additionally, the level of consumer vigilantism significantly influences the likelihood to share self-created misleading information with others. The unstandardized coefficient (0.26) is significant, $t = 6.61$, $p<0.01$, supporting Hypothesis 4. Consumers who act more vigilantly online are more likely to share self-created misleading information about a brand.

Mediation

The bootstrap method with 10,000 samples and a 95% confidence interval was used to calculate the mediating relationships (Preacher and Hayes, 2004; Hayes, 2018). The results of the mediation bootstrap analysis conducted in AMOS 24 are found in Table 2.5.

Dark Triad=> Proactive => Vigilantism

The a-path (0.14, $t = 3.93$, $p<0.01$) from the Dark Triad to proactivity and the b-path from proactivity to vigilantism (0.32, $t = 4.28$, $p< 0.01$) were significant. Consistent with Hypothesis

5, the indirect effect of the Dark Triad on vigilantism through proactivity toward opportunities, was also significant ($\alpha\beta = 0.05$, $p < 0.01$, CI Lower = 0.03, CI Upper = 0.10) (Zhao et al., 2010). Thus, Hypothesis 5 is supported; Proactivity toward opportunities mediates the relationship between the Dark Triad and consumer vigilantism.

Table 2.5 Hypothesis Results

Hypothesized Relationship		Estimates	t-Values	Direct Effect	Indirect Effect	CI: Low	CI: High	Hypothesis Supported?
Formative Indicators	H1a: Machiavellianism => Dark Triad	0.45	4.78**					Yes
	H1b: Narcissism => Dark Triad	0.34	4.54**					Yes
	H1c: Psychopathy => Dark Triad	0.18	2.08*					Yes
Structural	H2: Dark Triad => Proactive	0.14	3.93**					Yes
	H3: Proactive =>Vigilantism	0.32	4.28**					Yes
	H4: Vigilantism =>Share Self-created Misleading Information	0.26	6.61**					Yes
Mediation	H5: Dark Triad => Proactive => Vigilantism			0.28**	0.05**	0.03	0.10	Yes
	H6: Proactive =>Vigilantism =>Share Self-created Misleading Information			0.04	0.09**	0.04	0.16	Yes
	H7: Dark Triad =>Proactive=>Vigilantism =>Share Self-created Misleading Information			0.18**	0.01**	0.004	0.03	Yes
Index of Moderated Mediation	H8: Proactive * Vigilantism X Justify => Share Self-created Misleading Information				0.04**	0.01	0.10	Yes
	Vigilantism X Justify => Share Self-created Misleading Information	0.13	3.56**					
Interaction	Low Justify				0.04**	0.02	0.09	
	High Justify				0.13**	0.06	0.24	
Index of Moderated Mediation	H9: Dark Triad * Proactive * Vigilantism X Justify => Share Self-created Misleading Information				.001**	.000	.01	Yes
	Vigilantism X Justify => Share Information	0.09	2.74**					
Interaction	Low Justify				.003*	0.00	.01	
	High Justify				.01**	.003	.03	

Note: *indicates $p < 0.05$, **indicates $p < 0.01$, Model fit statistics: $\chi^2 = 1023.96$, $df = 441$, $p < 0.01$, $\chi^2/df = 2.33$, incremental fit index (IFI) = 0.92, Tucker-Lewis Index (TLI) = 0.91, comparative fit index (CFI) = 0.92, root mean square error of approximation (RMSEA) = 0.07
Model fit statistics are based on the full structural model, including the moderator and matched pairs latent interaction terms

Proactive=> Vigilantism => Sharing Self-created Misleading Information

A significant a-path from proactivity to vigilantism (0.32, $t = 4.28$, $p < 0.01$) and b-path from vigilantism to sharing the information (0.26, $t = 6.61$, $p < 0.01$) were both found. Consistent with Hypothesis 6, the indirect effect of proactivity on the likelihood of sharing the information through the mediator, vigilantism, was also significant ($axb = 0.09$, $p < 0.01$, CI Lower = 0.04, CI Upper = 0.16). (Zhao et al., 2010). Additionally, since the direct effect of proactivity in pursuing opportunities on the likelihood of sharing information is non-significant ($c = 0.04$, $t = 0.66$, $p = 0.51$), full, indirect-only mediation is found. Hypothesis 6 is supported; Vigilantism mediates the relationship between proactivity in pursuing opportunities and the likelihood of consumers sharing self-created misleading information.

Dark Triad => Proactive => Vigilantism=> Sharing Self-created Misleading Information

The a-path (0.14, $t = 3.93$, $p < 0.01$) from the Dark Triad to proactivity, d-path from proactivity to vigilantism (0.32, $t = 4.28$, $p < 0.01$), and b-path (0.26, $t = 6.61$, $p < 0.01$) from vigilantism to the likelihood of sharing the message were all significant. The indirect effect of the Dark Triad on the likelihood of sharing information through the mediators, proactivity, and vigilantism, was also significant ($axdx = 0.01$, $p < 0.01$, CI Lower = 0.004, CI Upper = 0.03) (Zhao et al., 2010; Hayes, 2018). Since the confidence interval does not pass through zero, the indirect effect is determined to be statistically significant (Hayes, 2018). Therefore, Hypothesis 7 is supported; proactivity and vigilantism drive the relationship between the Dark Triad and the likelihood of sharing self-created misleading information through a serially mediated relationship.

Moderated Mediation

The first step for using structural equation modeling to calculate moderated mediation, when using latent interaction terms, is to ensure that the full structural model has a good fit, which has been established (Cheung and Lau, 2017). Next, the index of moderated mediation is calculated to ensure the mediated effect is being moderated in the presence of the interaction term. Then the effect is reviewed by probing the interaction at high and low levels of the moderator (Cheung and Lau, 2017). The matched pairs approach is used to create latent interaction terms while minimizing any detrimental impact on model fit (Marsh, Hau, and Wen, 2004; Wu, Wen, Marsh, and Hau, 2013). Interaction terms are calculated to test the moderated mediation relationship (Frazier, Tix, and Barron, 2004; Spiller, Fitzsimons, Lynch, and McClelland, 2013). Before the interaction terms are calculated, standardized item scores are calculated to alleviate the risk of multicollinearity (Frazier et al., 2004).

Estimands are created to calculate the path coefficients at varying levels of the moderator –justification – at 1SD below the mean and 1SD above the mean to test for moderated mediation in the structural model (Bayl-Smith and Griffin, 2014; Maslowsky et al., 2015). These estimands conduct the bootstrap calculations while estimating the interaction effects created while allowing for multiple IVs (or a formative construct) (Bayl-Smith and Griffin, 2014; Little et al., 2012; Henseler and Chen, 2010; Thyroff and Kilbourne, 2018). The 10,000 bootstraps with a 95% confidence interval used in the mediation analysis were also employed when calculating the moderated mediation relationships (Hayes, 2018; Zhao et al., 2010).

Vigilantism and Justification

The mediated relationship that was found between proactivity and the likelihood of sharing information through vigilantism was reviewed once again by including the observed

justification in sharing the information. The interaction effect for Vigilantism \times Justification is significant (0.13, $t = 3.56$, $p < 0.01$). Additionally, the index of moderated mediation (0.04, $p < 0.01$) indicates that justification does moderate the full indirect effect on the likelihood of sharing the information. By probing the indirect effect of vigilantism on likelihood of sharing the information at various levels of the moderator at -1SD ($a \times b = 0.04$, $p < 0.01$, CI Lower = 0.02, CI Upper 0.09), and +1SD ($a \times b = 0.13$, $p < 0.01$, CI Lower = 0.06, CI Upper 0.24), a significant moderated mediation effect is found. Therefore, Hypothesis 8 is supported. As the consumer's justification of the decision increases, the relationship between proactivity and the likelihood of sharing the self-created misleading information through vigilantism is strengthened.

The Dark Triad, Vigilantism, and Justification

The mediated relationship that was found between the Dark Triad on the likelihood of sharing self-created misleading information through proactivity and vigilantism was reviewed once again by including the level of justification in sharing the information. The interaction effect for Vigilantism \times Justification is significant (0.09, $t = 2.74$, $p < 0.01$). Additionally, the index of moderated mediation (0.001, $p < 0.01$) indicates that justification does moderate the full indirect effect on the likelihood of sharing the self-created misleading information. By probing the indirect effect of vigilantism on the sharing the self-created misleading information at various levels of the moderator at -1SD ($a \times d \times b = 0.003$, $p < 0.05$, CI Lower = 0.00, CI Upper .01), and +1SD ($a \times d \times b = 0.01$, $p < 0.01$, CI Lower = 0.003, CI Upper 0.03), a significant moderated mediation effect is found. Therefore, Hypothesis 9 is supported. The consumer's justification in the decision significantly strengthens the relationship between the Dark Triad and the likelihood of sharing self-created misleading information through proactivity and vigilantism. The full model with results can be found in Figure 2.25

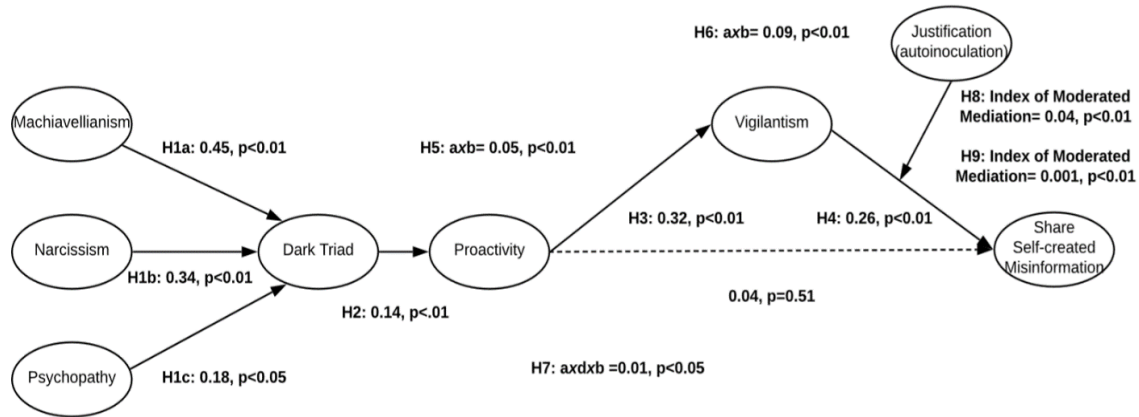


Figure 2.25 Full Model and Results

Discussion and Implications

The model, as a whole, describes how the Dark Triad collectively influences a consumer's proactivity in seeking out opportunities that benefit them, whether based on financial, amusement, entertainment, power-seeking or mischievous enjoyment motivations. Vigilantism in an online environment provides the proper channel for the consumers to facilitate that opportunity by using platforms to voice opinions and thus influence the likelihood of a self-created misleading message being shared. These automatic propensities are then strengthened by the autoinoculation that takes place through the justification for the creation and sharing of a misleading message. When this occurs, the consumer can build up immunity to not only counter views from others but also a response from a brand that may be the target of such maliciousness. Therefore, a consumer can then willingly create misleading information to use them to influence others.

This research contributes to the growing literature on the Dark Triad, inoculation theory, proactivity, and vigilantism by establishing the role of the constructs in driving the likelihood of

sharing misleading information. This study provides insights into the ways that online consumers can develop messages through both automatic and analytical means.

The Dark Triad is a formative construct comprised of Machiavellianism, Narcissism, and Psychopathy dimensions. This form of measure can be implemented in future studies to ensure that proper assumptions are maintained when measuring higher-order constructs. Advancing the distinctions and roles of Machiavellianism, Narcissism, and Psychopathy in forming the Dark Triad construct also provides an additional contribution by acknowledging that the formative model can provide a more appropriate means for measuring the Dark Triad variables as a whole. Since the formative construct captures the higher-order model, key distinctions can be made based on individual Dark Triad traits and the individual's overall propensity. This propensity can then result in the overall influence on the consumer's proactivity and vigilantism. Additionally, the role of all Dark Triad (Machiavellianism, Narcissism, and Psychopathy) as an overarching formative construct explains this relationship more effectively than each individual piece or a composite-based variable.

Dark Triad propensities drove the consumers' proactivity in pursuing opportunities, and this strongly influenced the level of vigilantism in influencing others. Additionally, these proactive consumers were likely to share misleading information when experiencing vigilantism propensities, as evidenced by the mediating relationships. The serial mediation effect indicated that individuals with higher Dark Triad levels are likely to be more proactive in getting what they want, using online platforms to voice opinions, and willing to share misleading information with others if it helps them to achieve their goals.

Another key finding is the role of justification in enhancing these actions. The study found that the mediated relationship became increasingly stronger in the presence of lower to

higher levels of justification. These findings suggest that propensities (or the automatic components of autoinoculation) are essential for action, but cognitive assessments through justification (analytical judgments) can be the deciding factor. These findings provide empirical evidence that the theorized Dark Triad traits interact with selfish motivations to encourage the sharing of misleading information with others when the consumer feels that their actions are justified.

Theoretical Implications

The studies provide numerous theoretical implications for understanding malicious consumer deviance in information sharing. First, the full implications and impact on inoculation theory revolve around the role of autoinoculation by the individual as a means to reinforce current views and actions. The uncovered role of autoinoculation extends inoculation theory as a dual-process model of decision-making and cognition. Consumers with negative proclivities are, therefore, more prone to employing both automatic and analytical cognitive resources to reinforce their own views. Moreover, when other consumers are deemed necessary to facilitate an opportunity, the originator of the information may act as the poster “zero” for the source of the informational inoculation and thus the source of group inoculation.

Second, this research explores the dynamics of conversations and messages that are shared with others in an online environment. This volume of two-way communication indicates that these individuals ultimately benefit from the sheer size of online environments in providing opportunities for entertainment, power-seeking, financial gains, and mischievous enjoyment. Inoculation theory, typically viewed from the standpoint of the marketer, may also be implemented by an individual consumer through autoinoculation (through the creation of misleading information) and the inoculation of others (through sharing the self-created

information). The consumers who seek to identify opportunities to benefit themselves may also need a critical mass of participation by others to achieve their goals. The findings indicate that the Dark Triad, proactivity, and vigilantism all contributed as a means to this end. However, the consumer was not likely to share purposely misleading information unless autoinoculation occurred in reinforcing the individual's propensities.

Third, this research extends the development of inoculation theory and autoinoculation, by illuminating the role of the Dark Triad propensities and proactivity, in driving vigilantism and misleading information sharing in online environments. This finding also highlights the online disinhibition effect since consumers are more likely to share purposely misleading information with others online than offline. These findings suggest that a crucial part of inoculation against contrary views is the initial development of motivations for disseminating the message in the first place.

Fourth, when the consumer experiences an opportunity to further their own goals, the automatic influence of the Dark Triad and proactivity in pursuing opportunities can drive the actions of a consumer in producing and sharing information. However, this automatic justification is enhanced by analytical cognitive justification. This finding indicates that individuals must exert mental resources through autoinoculation for the decision-making to occur. Thus, the studies suggest that inoculation theory can function as a dual-process model for decision-making that explains malicious information sharing and resultant behavior online.

Managerial Implications

The netnographic content analysis offers examples of the types of conversations that take place and the motives of consumers online. These propensities are then verified through quantitative analysis. Essentially, consumers develop implicit traits that drive behavior through

cognitive verification. When consumers seek to influence others in pursuit of an opportunity, their personal bias, beliefs, and justification can then drive their actions. The preceding studies provide at least four major implications for managers and practitioners.

First, the studies provide evidence that consumers with these propensities are quite active in an online setting and frequently create and communicate in communities to share both positive and negative information. What becomes difficult for practitioners is the ability to seek out and locate misleading information preemptively. As the posters communicate in close-knit settings, context, norms, and intent enhanced by shrouds of anonymity may be difficult to decipher by outsiders. Moreover, some messages may be sent in jest, while others may act as a call to action. Therefore, traditional tactics relying solely on automated processes for reviewing data may not provide the necessary contextual nuance to seek out and address the malicious information sharing. Companies must act quickly in determining which conversations are an active threat that should be addressed and which are harmless rants online.

Second, the Dark Triad traits occur in numerous consumer types. These propensities may vary in strength, but individuals who possess them are likely to seek out opportunities and utilize online platforms to influence others. Therefore, companies should be careful to monitor online consumer interactions when such information can be shared quite easily. This ease in information sharing is especially important since forums, message boards, and groups can be formed in real-time to exploit an opportunity. Managers should carefully monitor information sharing to offset any exploitative practices. Managers are also encouraged to scour through message discussions frequently, create baselines for individual posters, and build profiles for users who commonly begin discussions. Additionally, group conversations can be tracked through sentiment analyses or content analyses.

Although large organizations may be a larger target for such behavior (Kucuk, 2008; Kucuk, 2010) and have more flexibility in assigning resources, small-and-medium-sized companies can still set standards for social media specialists and managers to review and assess online misleading information sharing. Moreover, companies can begin to analyze the key discussion leaders and other topics that are addressed to determine if certain participants exert more credibility than others. In doing so, a manager may uncover potential themes that can quickly lead to coordinated misleading information if an opportunity presents itself.

Third, based on the findings of the study, the individual sharer starts the process of inoculating other consumers against counter-arguments from brands. For instance, a consumer who autoinoculates through the justification of actions can protect his or her views from scrutiny by vocalizing the justification to others. Undoing this process can be daunting since the consumer may have a bias confirmed through the autoinoculation. A brand may have difficulty in persuading the individual to change their viewpoint. Therefore, the influence of the original poster and the group can lead to much broader implications through larger-scale harmful information sharing. When a consumer begins sharing negative misleading information with others that confirms a belief, a firestorm, or sudden discharge of messages containing negative information against the brand, may begin (Pfeffer, Zorbach, and Carley, 2014; Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). Because of these daunting tasks, firms can benefit from diligent “re-information” or “pre-information” campaigns to make sure that those that may be prone to believing the misleading information are not fully susceptible to the impact of the inoculation from the poster. These campaigns could include participation in similar posts, adding to the conversation, or reinforcing the brand image through integrated marketing communication.

Finally, the true damage of one poster comes from the multiplier effect of misleading information through group sharing. If managers can work diligently to build positive word of mouth and reinforce brand image, the original poster may have difficulty in swaying the views of others. By countering the availability of negative information with positive information, bystanders exposed to the misleading post may choose not to join in the information sharing since other posters do not corroborate the information. Companies that attract consumers online should be diligent in monitoring information while tactfully encouraging and empowering pro-brand groups and brand communities to act as advocates. In doing so, companies can track the actions and information shared by individuals online to determine which consumers may be at risk of sharing misleading information. Moreover, brands can act to empower brand communities to seek out at-risk communities for malicious information sharing before issues arise.

Limitations and Future Research

The netnographic content analysis focused only on observations without researcher participation in the forum. Future studies could actively pursue additional understandings by encouraging researcher participation and interaction with the posters. Another limitation is that the qualitative study focused exclusively on Reddit forums. Although Reddit forums provide an appropriate sample of online communication, future studies should examine other social media platforms. Additionally, other methods, such as web crawlers or sentiment analyses, can be implemented to retrieve rich data insights from the forums. An M-Turk sample was used in the pretest of the survey items, and an M-Turk sample was used to collect data in the primary study. Since the study focused on whether or not a shopping situation took place online or offline, the sampling method was deemed appropriate as a wide range of consumers participate in both shopping situations. Also, the author set parameters that would only include U.S. respondents,

and only those with a 95% work acceptance rate were considered for participation. While care was taken to provide a scenario appropriate for these samples, the methods employed best practices for both the student snowball sample (Yoo and Donthu, 2001; Petersen and Marunka, 2014) and the methods employed best practices for M-Turk (Hulland and Miller, 2018; Hulland, Baumgartner, and Smith, 2018; Sharpe Wessling, Huber, and Netzer, 2017), future research might replicate these findings using a different sampling approach. Additional scenarios could be used to triangulate the results and further differentiate the various online situations and determine the role that these consumer types have on others. The next studies will also uncover the role of consumer communities in sharing information, which messages are most influential, and the role of critical mass in message sharing.

Conclusion

The purpose of this research is to start a larger narrative in understanding misleading online information sharing and the significant role that individuals play. Online consumer interactions are vastly complicated and should be studied diligently to understand the nuances of individuals. When these tendencies are examined, a larger picture can be observed to understand the role of individuals in the larger paradigm of misleading information sharing. The automatic and analytic components of decision-making provide an individual with motive and intent to pursue their desires. When this drive is enacted toward good, other consumers and companies can benefit. However, when enacted for malicious intent, or the confirmation of biases, other consumers may be misled.

CHAPTER III

ESSAY TWO

The Herd and the Bystander: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments

Abstract

Consumers with malicious motives can aim to destroy brands by posting misleading information online. However, without participation of other consumers, these initial messages may go unnoticed and fade away. When a message breaks through the noise online, herd immunity can develop to sway larger groups and reinforce the views of the malicious poster. In these situations, a cascade of negative information can act as a mechanism for collective belief formation that initiates a faux storm of misinformation. This confirmation through mere exposure can then flow to other consumers and develop further credibility. Fortunately, a brand community response can act as a barrier to help offset the impact of a Negative Cascade, but this may be easier said than done. The implicit bystander effect states that an individual may be likely to relegate actions to others, especially in the presence of larger groups. Therefore, the individual community or forum members are unlikely to offset the negative information completely nor to post positive messages since they may feel that other members will share posts to correct the erroneous information. This essay will explore the types of messages and the strength of a Positive or Negative Cascade in driving the dissemination of misleading information and the challenges of cultivating an effective community response.

“What's really funny is that early on in the post, roughly an hour in, someone posted to please cancel orders as it was due to a larger distributor error and that someone would get the axe for such a mistake, and instead they [that poster's comment] got downvoted into oblivion”

2

Introduction

Individuals posting fake or fraudulent information erode trust (Barnett, 2014; Ognyanova, 2019), spearhead negative information sharing (Appel, Grewal, Hadi, and Stephen, 2019), and further confuse consumers (Clark, 2019; Fitzgerald et al., 2019). As found in Essay 1, these individuals exhibit tendencies toward the Dark Triad (Machiavellianism – manipulation, Narcissism – self-grandiosity, and Psychopathy – callousness) (Buckels et al. 2014). Moreover, these consumers are able to seek out opportunities to benefit themselves online and do so by engaging in vigilante tendencies by using the forum or community as a soapbox for their ulterior motives. In doing so, these consumers aim to sway both other consumers and brands in actions that are taken to enhance themselves.

Onlooking consumers, or bystanders, may accept a false message about a brand unwittingly, with minimal counter-argument since the pursuit of new knowledge could be time-consuming, and the current information “seems” accurate, especially if prevailing views are confirmed. In isolation, the effects of this misinformation can dissipate. However, when false

² The quote from above is from an online consumer on Reddit describing his/her reaction to misleading information shared by other posters about an online pricing mistake. Accurate information that was provided to counter misleading information, shared by those hoping to get the deal, was downvoted, making it less likely to be seen by most participants in the group.

messages are disseminated to a group – as is often the case in these situations –, herd inoculation, group-level belief, and support that diminishes external information effectiveness, can begin to take effect since numerous sources create additional points of message confirmation (Banas and Rains, 2010; Miedema, 2018). Consumers may then assume that an innocent brand is guilty of a transgression until they receive sufficient additional information to prove otherwise. This phenomenon becomes daunting for brands to address since larger scale negative information sharing becomes increasingly difficult to mitigate (East, Hammond, and Wright, 2007), reputations can be destroyed in hours (Horvath, 2014; Thomas, 2013), and bias in the information is likely present (Peterson and Merino, 2003).

Consumers bombarded with stimuli from numerous sources throughout the day cannot possibly process every minute detail. Attempting to minimize the effort required to handle so much information, they may use decision-making aids (Kahneman, 2011; Kahneman and Frederick, 2002). These aids may include reviews, references to prior experiences, or available information from other consumers. However, customers overloaded with information may make bad decisions (Levetin, 2016). Moreover, nearly 43% of social media users have *shared* fake or misleading information (Woodcock, 2019). In these situations, consumer bystanders, those who happen to be present when false information sharing is taking place, may then look to the other group members to determine whether they are (Fischer et al., 2011) whether by condoning the information or challenging it (Garcia, Weaver, Darley, and Spence, 2009). The quality of information can be difficult to assess, and the ulterior motives of others may go unnoticed. When information is received in an online environment, processing, decoding, and discriminating factual from un-factual (purposely-incorrect) information can become a daunting task. In such

situations, consumers may resort to emotional appeals that are vivid and passionate rather than factual information that relies on slow and difficult deliberation (Roberts, 2016).

Understanding how misinformation from an original poster spreads throughout the herd of a broader consumer ecosystem online is paramount to the creation of an effective response to such actions, especially when consumer bystanders generate further and continual negative belief as additional contemptuous information is shared (Johnson, Matear, and Thompson, 2011). The concept of availability cascades, collective beliefs that develop through repeated information exposure (Kane and Webster, 2010), explains the mechanisms at play in driving this phenomenon by addressing the broader role that a community plays in information dissemination through herd inoculation. The true power of an availability cascade is the sheer abundance of information that provides credence to the message. If a prevailing sentiment permeates throughout a discussion and becomes the bulk of available information, then these messages become more convincing to others. Such as in Reddit forums, the loudest voices, those posts that generate active discussion and up/downvoting from others, are typically seen as the most accurate or truthful.

Research on how messages spread online has typically focused on an initial dissatisfaction such as a service failure driving information sharing within consumer communities (c.f. Larson and Denton, 2014; Daunt and Harris, 2011; Kucuk, 2010; Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). However, a purposely-misleading message may be seen as authentic and can be accepted as factual by consumers. Since malicious posters can utilize forums in real-time to engage others, this topic is necessary for determining how the

initial message leads to a much larger brand crisis through the creation of a faux storm of misinformation.

While researchers have addressed the role of online communities in consumer responses to actual service failures, little research has examined the role of an individual and the community in sparking a brand crisis in the absence of a real failure. Although studies have reviewed the actions of the poster (Rauschnabel, Kammerlander, and Ivens, 2016), message tones that are most effective (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019), and the associated content that reinforces a message (Hewett, Rand, Rust, and van Heerde, 2016). Little research discusses the mechanisms that drive the false message sharing in larger and diverse consumer ecosystems online. Such an understanding is imperative, since a Negative Cascade of information can enhance message belief. Very little information is available to describe the types of messages that are most believable, and the concept of an availability cascade as the mechanism driving the spread of information is underexplored. Therefore, this essay aims to answer the overarching question. “How does misleading information spread through online communities?”

The studies that follow contribute to the understanding of how the believability of the original post is enhanced by an associated cascade of information that influences the likelihood of a message being continually shared by other consumers. Inoculation theory (Banas and Rains, 2010) provides the theoretical lens in describing the role of a larger group of consumers in spreading inoculation and enhancing herd immunity to either halt or continue the spread of misinformation. The findings suggest that the initial post can be seen as believable when the message conveys an emotional tone while using wording that conveys authenticity.

Additionally, the forum itself may act as fuel that enhances or detracts from continual sharing of the misinformation. When a negative message is paired with a Negative Cascade, a truth default effect (Levine, 2014) can drive future sharing since the consistency makes consumers assume that the truth was communicated by both the original poster and corresponding forum. Whereas, a negative message that is paired with a Positive Cascade that defends the brand can break the continued information sharing and discredit the original poster. Therefore, the herd inoculation effect is enhanced in the presence of a negative confirmation cascade of information, and a Positive Cascade reverses the herd inoculation that occurs. The bystander effect arises more so when consumers are exposed to conflicting information and choose not to involve themselves in the discussion. This conflict in information requires both automatic and analytic resources for determining if actions are necessary. When the choice to act is difficult or other parties appear to be defending a maligned brand, a bystander may choose to deflect positive or negative posts to others in the community.

The theoretical framework is discussed, then the methods are explained, followed by a description and discussion of the positive and negative availability cascade results. The cascade results are tested again to determine the presence of the bystander effect in community information sharing.

Theoretical Framework

Inoculation Theory: A Dual-Process Model

Inoculation theory describes the way that individuals can thwart persuasive attempts from other parties (McGuire, 1968). Dual-process models of decision-making (Kahneman, 2011; Kahneman and Frederick, 2002) propose that our ability to produce thought and make decisions comes from both an automatic response system (influenced by bias and heuristics) and a more

alert cognitive system that requires more effort to implement (Kahneman, 2011). Inoculation theory suggests that arguments that confirm a current belief are difficult to offset. Additionally, individuals who have been inoculated against a message can be tough to persuade when conveying a counterargument as more confirmatory data points become available (Lessne and Didow, 1987). When a large group reinforces the original claim of an individual, through herd inoculation, much more cognitive effort is required to offset its effects. Although inoculation is typically seen as defense against an external threat (i.e. from the brand or consumer's perspective), the herd inoculation developed in a forum or community is within members to counteract the threat of a brand response or other consumer response that challenges the shared negative view initiated by the masked agenda of the original poster. Therefore, inoculation can reinforce misinformation in a similar manner to factual information.

Herd Inoculation

A consumer who is told that a brand is uncaring or unethical may adopt this view unless ample evidence is provided to the contrary. However, supplying this necessary evidence may be difficult, when large amounts of negative information confirm this prior belief. This theoretical framework describes the real strength of the Negative Cascade in reinforcing the inoculation through repeated but unobvious repetition of exposure to misinformation (Banas and Rains, 2010). Moreover, this repeated exposure by others enhances the inoculation of others by creating a herd immunity to counter persuasion.

Dual-process models explain the influence of online reviews and existing opinions on other consumers' quality perceptions (Filiberti, 2014), initial impressions of others (Ingold, Donni, and Lievens, 2018), and perceptions of employer brands (Rampl, Opitz, Welp, and Kenning, 2016), but they have not been explored as the mechanisms driving malicious and misleading

information sharing. Automatic responses, such as a gut reaction, can dictate the initial evaluation of information. When the information is readily available, the inoculation process can be enhanced. Moreover, the use of more cognitive resources can be daunting for consumers, and accessing other information may be perceived to contribute to only marginal returns in knowledge. Therefore, automatic responses are employed to minimize the use of mental resources (i.e., working memory) (Evans and Stanovich, 2013). Thus, unless an individual is otherwise motivated to employ additional resources to question the message and increase involvement, they tend to resort to prior assumptions (Pfau, 1997). Essentially, without putting in additional effort to strengthen a new view, the consumer will opt to confirm a bias. If the only information available is confirmatory, then it can enhance the herd immunity of the group as the information spreads.

For instance, individuals may see a negative post and then find a follow-up discussion that confirms its content. These consumers may then immediately assume the original message was factual and determine that additional data is not necessary. The bystander effect suggests that consumers are less likely to come to the defense of another – a brand in this instance – when they are in a group setting – an online forum in this case. Additionally, the same can apply to positive information. A diffusion of responsibility reduces the chance that they will confront the negative tone. The original post may not have its intended effect since the conflicting positive information reduces the likelihood of belief. However, this positive information in a Positive Cascade does not guarantee action by the individual since the conflicting information requires greater effort in decision-making.

The role of inoculation theory as a dual-process model of decision-making provides the appropriate lens to understand both the role of availability cascades in driving the dissemination

of misleading information and the role of the bystander effect in limiting the response brand or other consumers.

Availability Cascades

Since individuals utilize the availability, heuristic in decision-making, information that is seen often is deemed to be true (Kuran and Sunstein, 1998). An availability cascades develops when a collective belief gains plausibility as it is continually shared throughout online communication (Kane and Webster, 2012). When forum members share a number of messages conveying a belief or view, other consumers can be exposed to this information on additional platforms. For example, a fringe website may go unnoticed until topics are shared on Reddit which then spreads to Google results that then make the views more readily accessible and more plausible. Availability cascades influence the development of belief in others tied to social or health issues (Barr, 2013), confirmation bias (Alfano, 2011), political views (Sun, 2012), corporate regulation (Kuran and Sunstein, 1998), and information credibility (Gaziel Yablowitz and Raban, 2016). When larger and larger groups of consumers reinforce information-sharing, the original message is heavily reinforced, much like group-level or herd inoculation. Although a cascade of information can sometimes be positive, it is the information sharing that takes on a negative tone that causes harm to brands.

Further, firestorms are defined as “a sudden discharge of large quantities of messages containing negative WOM against a company” (Pfeffer, Zorbach, and Carley, 2014 pg.118) that may result as a byproduct of a Negative Cascade. When a firestorm occurs, a brand must diligently act to diffuse it since the lack of a response can exasperate the harmful information sharing (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). Close-knit communities can

act to reinforce the message (Hewett, Rand, Rust, and van Heerde, 2016) and may thus enhance herd immunity. Those who are engaging in online malicious behavior benefit from the organic formation of a Negative Cascade that flows through discourse to others and creates a faux storm. When other consumers begin sharing misleading information, that exacerbates the faux storm, the original poster may have little need for further action. A Negative Cascade of misinformation can occur because the initial misleading message empowers others to share additional opinions or beliefs that may or may not be legitimate. When a consumer communicates in an online setting, many pieces of information can reinforce a current view and distort perceptions of others (Kuran and Sunstein, 1998). A consumer may be more likely to believe the fraudulent information provided by a poster if the message is frequently posted and readily available (Pollock, Rindova, and Maggitti, 2008).

Since multiple sources can share online information, the illusion of truth may be present as each account appears to be “first-hand,” and thus a firestorm of information develops around the shared belief (Dechêne, Stahl, Hansen, Wänke, 2010; Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). As negative information becomes more available, other posters involved with the message can become more vested in information sharing (Koh, Gunasekaran, and Rajkumar, 2008). These individuals who have been inoculated may actively preempt corrective brand response to reduce the likelihood that others change their position. Therefore, group herd inoculation may become even stronger when the Negative Cascade acts as the mechanism for confirming the misinformation. However, a brand may also utilize the same mechanisms by disconfirming a message through a Positive Cascade that contradicts the original post.

Believability of the Message

For an individual to be inoculated, they must first be exposed to a preemptive threat or initiating message. Initial belief in this content sets the stage for rumination and acceptance. An initial narrative can provide a subtle nudge by specifying an event, conveying a concern, or sharing a negative view with benign intent. The believability of the message, a view that information is true and acceptable, enhances product evaluations (Gürhan-Canli and Maheswaran, 2000), consumer coordination (Sen, Sankar, Zeynep Gurhan-Canli, and Vicki Morwitz, 2001), and increases the trustworthiness of a source (Moore, Mowen, and Reardon 1994). However, individual motives can drive deviance, warp the intent of the message, distort the content, and mask intentions of the sender (Jonason, Webster, and Schmitt, 2009). Consumer bystanders may pay close attention to the tone of the message being conveyed to determine credibility and the legitimacy of the message.

The emotional tone of a message has been shown to influence the belief in negative WOM when the overall tone is negative (Herhausen et al. 2019). However, a negative tone in itself will not necessarily enhance the believability of content. Messages that convey social, personal, and ethical experiences further enhance the believability of a message (Schwarz, 2004). By communicating information effectively, the sender can capitalize on truth default assumptions of the receiver (Levine, 2014). Accordingly, when the negative tone of a message is matched with a description of the event with context, it can drive a greater belief in the source. For example, an angry and belligerent poster online may not have the same impact in convincing others as someone posting negative information in a more calculated and descriptive manner since narratives influence immersion in the content (Kahneman, 2011). Because a negative emotional tone paired with situational context has been shown to influence believability, it is

expected that the tone of the message from the original poster will influence perceived believability by the recipient bystander. Therefore:

Hypothesis 1: A contextualized message from the original poster will positively influence a consumer's belief of the message than a message of purely negative emotional tone alone.

Justification

When an individual justifies the feelings felt or motivations realized, s/he feels more comfortable with taking action through information sharing. Justification in information sharing suggests that the message is deemed acceptable and defensible (Inman and Zeelenberg, 2002). A consumer who can justify decisions and outcomes also reduces dissonance and reinforces outcomes experienced. Justifying a message is a fundamental component of inoculation in that the receiver must assess the information and determine its legitimacy. Moreover, justification requires cognitive resources (Kim, Kim, and Park, 2012), and implementing these resources can be influenced by the presence of new information (Park and Cho, 2012; Kline and Wagner, 1994).

When information is readily available to help form an opinion, this justification can be motivated by more impulse-based decision-making (Spears, 2006). When the message received is deemed acceptable, the consumer bystander may then feel justified in sharing that message with other parties. When the initial "gut" reaction is to believe a statement, the individual may resort to heuristics, such as the content (Meyers-Levy and Maheswaran, 2004), tone (Werle and Cuny, 2012), or quantity of a message (Dawson and Brashers, 1996). Additionally, when a consumer deems that the actions of another are justifiable, those actions are more likely to be endorsed (Malaviya and Sivakumar, 2002).

The justification for sharing a message must be present for the receiver to warrant the decision to take additional action. When the content of information confirms an initial availability heuristic, even through one data point, justification for sharing the message can be more easily explained to oneself and others (Tversky and Kahneman, 1973). Thus, when a message is deemed believable, it is therefore deemed as more justifiable to be shared. Therefore:

Hypothesis 2: The consumer's belief in the original message will positively influence the perceived justification of spreading the message

Likelihood to Spread the Message

Numerous outlets are available where posters can post and share content (i.e., Reddit, Facebook, Twitter). Consumers can quickly determine whether a message is worth sharing based on their personal interest in the topic (Chalkiti and Sigala, 2008), negativity (Chen and Lurie, 2013), or frequency (Kahneman, 2011). When the messages are readily accessible, consumers can readily be exposed to hundreds of messages per day with the option to then share these stories with others. Consumers can easily choose to click a button and share a message with their contacts. However, not all consumers post negative brand information on social media. Just because a message is believable does not mean that consumers will share that content with others. The consumer must justify the sharing of the message. This key threshold of justification in sharing the message solidifies the groundwork for actions that are taken (Malaviya and Sivakumar, 2002). Consumers further develop social connections and aim to inform or persuade others (Gatignon and Robertson, 1985). Therefore, the justification in sharing the message with others, whether it be to inform friends (Dubois, Bonezzi, and De Angelis, 2016), discuss a risk (Akpınar, Verlegh, and Smidts, 2018), or warn others (Blodgett, Wakefield, and Barnes, 1995) can enhance the likelihood that a message is shared.

Consumers may also be more likely to share a message that confirms a particular belief or preference (Lallement, Dejean, Euzéby, and Martinez, 2019). This confirmation bias suggests that information that is shared online may not provide evidence that an event actually occurred but rather suggests that the individuals receiving the message see it is probable. When this occurs, the individual is more likely to share the original message with others online. Therefore:

Hypothesis 3: The consumer's perceived justification for spreading the message will positively influence their likelihood to spread the message.

Believability and Justification

The original message viewed online can begin laying the groundwork of the automatic components of inoculation. An individual's automatic propensities provide a starting point for actions and simple decision-making (Kahneman, 2011). A gut reaction, heuristic, or individual characteristic can drive the types of actions pursued since fewer cognitive resources are required to act (Evans and Stanovich, 2013). Once the message is viewed, a consumer may claim the message "looks legitimate." Thus, consumers exposed to a negative message from a poster become important influencers of future message sharing when a mob mentality or herd immunity begins to develop (Johnson, Badger, Waltermire, Snyder, and Skorupka, 2016). Messages that are deemed more believable help to streamline this process by providing legitimacy and trustworthiness to the content (Munnukka, Uusitalo, and Toivonen, 2016).

Additionally, the believability of a message is necessary but not sufficient to drive the sharing of content. When a consumer accepts the content as more factual, s/he the justification in sharing a message is also enhanced. When the consumer experiences cognitive justification in sharing a message, then the inoculation initiated by the original poster can be continually spread

to other consumers. The newly inoculated consumer then continues the inoculation process by enhancing herd inoculation through spreading the message and increasing availability.

Since the relationship between the original message content and likelihood to spread the message is expected to flow through the believability and justification in information sharing, it is also predicted that believability and justification will mediate the relationship between the original poster's message and the likelihood of spreading the message. Therefore:

Hypothesis 4: The consumer's belief in the message and perceived justification in sharing the message will serially mediate the relationship between the original message and the likelihood to share the message.

Negative and Positive Cascades

The original message may go unnoticed unless other posters get involved in reinforcing the content. A cascade can reinforce the legitimacy of a message online by implying greater trustworthiness of the message (Kane and Webster, 2012). Additionally, the presence of a cascade of confirmation information may allow a consumer to resort to more heuristic or automatic processing since the volume of information can reinforce the original message (Kahneman, 2011). Thus, it is expected to influence the inoculation of the consumer bystander since the automatic responses dictate the evaluation of information credibility through availability and enhance the inoculation process.

Moreover, the use of more cognitive resources can be daunting for consumers, and accessing additional information may be perceived to contribute to only marginal returns in knowledge. Additionally, a cascade of information that contradicts the original poster will halt the believability of the message by providing access to counterinformation in greater quantity.

The mere availability of negative information can ease the process of decision-making and influence the likelihood that the original poster's message is shared. Thus, the presence of an availability cascade is expected to strengthen the mediated relationship between the original poster's message and the likelihood of spreading the message belief and justification in message sharing. However, the availability of positive information (that disconfirms the original message) causes the consumer to question the original message and is expected to weaken the mediated effect between the original poster's message and the likelihood of spreading the message belief and justification in message sharing. Therefore, an interaction between the original authentic message and cascades of negative (confirmation) or positive (disconfirmation) information is expected. Therefore:

Hypothesis 5a: The presence of a Negative Cascade of confirmation information will strengthen the serially mediated relationship between the original message and likelihood to spread the message through the consumer's belief in the message and justification of the message.

Hypothesis 5b: The presence of a Positive Cascade of disconfirmation information will weaken the serially mediated relationship between the original message and likelihood to spread the message through the consumer's belief in the message and justification of the message.

See the full conceptual models for Studies 1a and 1b in Figure 3.1 below.

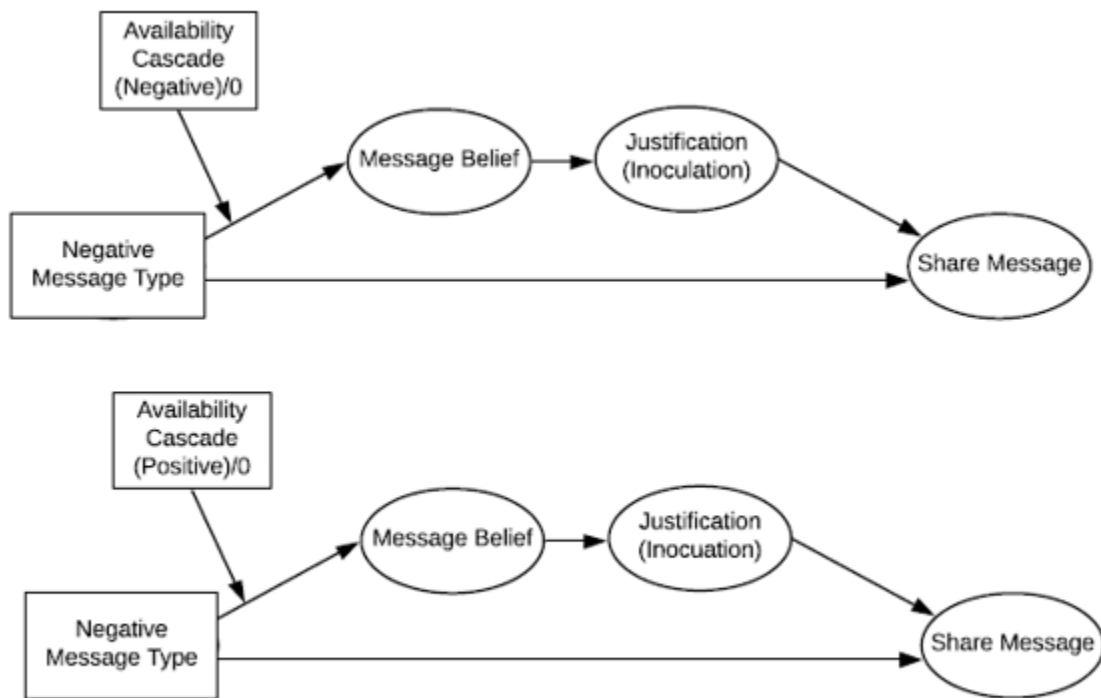


Figure 3.1 Conceptual Models: Studies 2a and 2b

Study 1: Message Type

Message Type Believability

An overarching negative tone in a forum is a sign that a firestorm could be taking place (Herhausen et al. 2019). However, messages from individuals also have other components beyond just raw negative emotion. Therefore, additional context is needed to determine the type of message most likely to initiate a faux storm online. Message types are created using Linguistic Inquiry and Word Count (LIWC) to convey a negative tone with corresponding authenticity, clout, and authenticity. A MANOVA is conducted to determine the most effective message type on believability, justification, and likelihood to spread the message as a whole.

The respondents are exposed to one of the four LIWC treatments (Analytical, Clout, Authentic, or Negative Emotional Tone) that discusses a posted claim that New Latitudes deletes negative reviews. The scales used to measure the believability of the message (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000), justification of the message (Inman and Zeelenberg, 2002), and likelihood to spread the message (Lee and Ma, 2012) were adapted from existing research scales. All scale items used a 7-point Likert-type or semantic differential endpoints. This knowledge is essential for the upcoming studies since the initial narrative can be crucial in deciding the believability of the message and the corresponding information sharing.

LIWC Treatments

To test the most believable message type, the author developed negative treatments based on the summary variables from LIWC (Linguistic Inquiry and Word Count) that correspond to various messages found in an online environment discussing a brand in a negative light. The four summary variables are discussed below.

The analytical thinking variable is a factor-analytically derived dimension based on eight cognitive-function word dimensions (c.f. Pennebaker, Chung, Frazee, Lavergne, and Beaver 2014). The analytic dimension addresses the degree to which people use words that suggest formal, logical, and hierarchical thinking patterns (Pennebaker, 2019; Pennebaker, Chung, Frazee, Lavergne, and Beaver, 2014). High scores on analytical thinking tend to use language in less narrative ways and focus on facts rather than experiences.

The clout variable refers to the relative social status, confidence, or leadership that people display through their writing (Pennebaker, 2019). The algorithm was developed based on the results from a series of studies where people interact with one another (Kacewicz, Pennebaker,

Davis, Jeon, and Graesser, 2013). Individuals may convey a certain level of importance or status based on their postings. This variable provides insights into an individual's attempt to sway others.

Posts that appear to be authentic can reinforce the message that is being conveyed. Newman, Pennebaker, Berry, and Richards (2003) developed the algorithm for authenticity through a series of studies in which people were convinced to be honest or deceptive. Additionally, Pennebaker (2011) contributed to the creation of the algorithm through published summaries of deception studies. When people reveal themselves authentically or honestly, they are perceived as more personal, humble, and vulnerable (Pennebaker, 2011). Additionally, posters may convey information in a way that is deemed more authentic to influence the other members of the online group.

Finally, the emotional tones are determined for both the positive and negative emotion dimensions (Cohn, Mehl, and Pennebaker, 2004). Pennebaker (2011) states that the algorithm is built so that the higher the number (i.e., 100), the more positive the tone. Scores below (50) indicate a more negative emotional tone (Cohn, Mehl, and Pennebaker, 2004). This measure addresses the raw emotion that is conveyed in posts. Since a negative forum tone is common in firestorms, the treatments each contain a negative emotional tone along with the corresponding analytical, clout, or authenticity treatment.

MANOVA

To test for the most believable message type, the author developed negative treatments based on the summary variables from LIWC (Linguistic Inquiry and Word Count) that correspond to various messages found in an online environment discussing a brand in a negative light. The participants were exposed to one of the dimensions (Analytical, Clout, Authenticity,

Negative Emotional Tone) created using LIWC (see Figure 3.2). Since LIWC provides the summary variables based on the message content, the researcher used the tool to generate a message high in each dimension.

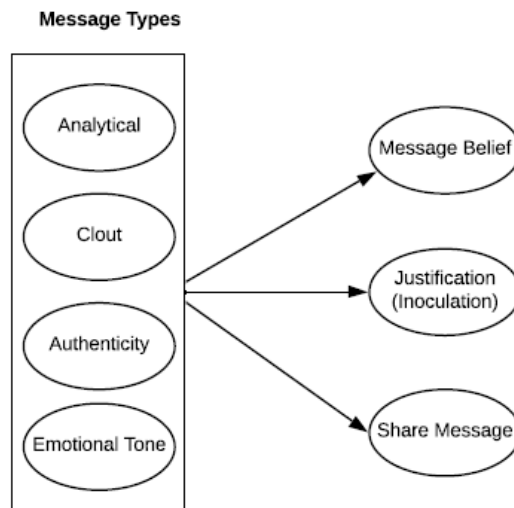


Figure 3.2 MANOVA

Procedure

A total of 251 participants that adequately represent online shoppers were recruited from an M-Turk panel. Within the sample, 51.4% were male, and the average age of respondents was 36 years old (see Table 3.1). M-Turk was chosen since respondents represent online consumers and appropriate sampling frame and can produce reliable and consistent results (Goodman et al., 2013). Respondents were shown an image and description of an online company called New Latitudes that sell products through online channels such as Amazon, Etsy, and their own websites. To ensure quality online respondents, various attention checks were required to ensure respondents were not selecting random answers, respondent screening policies that required a

95% HIT success rate, and a requirement of written text for each respondent to describe what they saw were included in the study. Additionally, respondents were required to rate each treatment based on the assigned LIWC treatment. Thought listings that required participant input in the survey were used to screen responses and were reviewed for any abnormalities, and no apparent problems were evident.

The total responses (251) were exposed to a negative emotional tone paired with the LIWC summary variables. These treatments included 64 participants exposed to the Analytical treatment that conveys an overall negative tone using words that convey cognitive processes and analysis, 61 exposed to the Clout treatment which conveys social status and expertise, 63 exposed to the Authenticity treatment, and 63 exposed to the Negative Emotional Tone Treatment alone.

Table 3.1 Study 1 Sample Description

Category	MANOVA
Age	
18-20	7 (3%)
21-30	89 (35%)
31-40	83 (33%)
41-50	47 (19%)
51-60	18 (7%)
61 and over	7 (3%)
Gender	
Male	129 (51.4%)
Female	122 (48.6%)
Social Media Usage	
Never	12 (5%)
Monthly	13 (5%)
Weekly	34 (14%)
Daily	157 (63%)
Hourly	35 (14%)

New Table 3.1 (continued)

Social Media Discussion Activity	
Not very active	38 (15%)
Not active	53 (21%)
Neither active nor inactive	51 (20%)
Active	88 (35%)
Very active	21 (8%)

The reliability of the scale items was assessed, and each exhibited an acceptable level of reliability ($\alpha \geq .80$, Nunnally and Bernstein 1994). A confirmatory factor analysis was then performed using AMOS 26 to assess the unidimensionality, convergent validity, and discriminant validity of the latent constructs. The results of the analysis indicated an acceptable fit (Marsh, Hau, and Wen, 2004) of the model to the data ($\chi^2=116.71$, $df = 51$, $CFI = .98$, $TLI = .97$, $IFI = .98$, $RMSEA = .07$). Table 3.2 shows a complete list of results from the CFA, along with composite reliability for each construct. The convergent and discriminant validity of the scales was examined by calculating the average variance extracted for each construct. Furthermore, calculations for the shared variance between constructs were made as recommended by Fornell and Larcker (1981). The average variance extracted was at least .50 for each construct, providing evidence of convergent validity. No shared variance measures between constructs exceeded the average variance extracted, providing evidence of discriminant validity. Common method bias was then assessed since the data were collected at a single point in time. A single latent construct was included in the analysis to control for any bias resulting from the common method (Podsakoff, MacKenzie, Lee, and Podsakoff 2003). This factor included a specified relationship to each scale item to account for any systematic bias within the latent constructs.

Table 3.2 CFA Results, t-values, and Cronbach's α

Scales	α	Standardized Factor Loadings	t-values
Believability – (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000)	0.93		
Highly believable		0.75	15.29
Absolutely true		0.82	18.14
Totally acceptable		0.65	12.06
Very credible		0.87	20.55
Completely trustworthy		0.92	*
The justifiability of the decision - (adapted from Inman and Zeelenberg, 2002)	0.90		
Strongly Justifiable		0.93	29.49
Very Logical		0.92	27.67
Acceptable		0.91	28.77
Reasonable		0.95	*
Likelihood to Spread the Message (Lee and Ma, 2012)	0.95		
I would share the message online		0.94	34.27
I would share the message when discussing the company		0.90	24.94
I would spread the information to others		0.99	*

Model fit statistics

Note: $\chi^2 = 116.71$, $df = 51$, $p < 0.01$, $\chi^2/df = 2.29$, incremental fit index (IFI) = 0.98, Tucker-Lewis Index (TLI) = 0.97, comparative fit index (CFI) = 0.98, root mean square error of approximation (RMSEA) = 0.07.

All factor loadings have a p-value < 0.01

* denotes a constrained relationship to 1.00 for identification

As an additional test of discriminant validity, correlations between the individual variables were compared to the square root of average variance extracted (Fornell and Larcker 1981). As illustrated in Table 3.3, which also displays the means, standard deviations, and correlations between variables, the square root of AVE for each variable exceeds the variance that each variable shares with the other variables measured in this study.

Table 3.3 Means, Standard Deviations, AVEs, and Shared Variance

	Mean	SD	1	2	3
Believability	4.40	1.21	0.65		
Justifiability of the Decision	4.25	1.79	0.50	0.86	
Likelihood to Spread the Message	3.37	1.93	0.33	0.52	0.86

Note: Average Variance Extracted in the diagonals

To test if there were meaningful differences among the four different LIWC summary variables, the total effect on all constructs was compared using multivariate analysis of variance (MANOVA) on the believability of the message, justification of the message, and likelihood to spread the message. MANOVA was considered an appropriate method for this analysis as it allows for analysis of multiple dependent variables simultaneously, with sufficient power (Hair et al. 2010).

Results

The results of the MANOVA did not show a statistically significant difference between the four LIWC treatments on the believability of the message, justification of the message, and likelihood to spread the message ($F(9, 596) = 1.56$; Wilks Lambda=0.94, $p = .09$). To find if variables contributed to the significant differences between summary variables, an ANOVA was conducted to view each variable individually. The results from the ANOVA are shown in Table 3.4

Table 3.4 ANOVA Results

Variables	Analytical	Clout	Authenticity	Negative Tone
Believability	4.50	4.29*	4.76**	4.12**
Justification	4.43	4.08	4.43	4.06
Likelihood to Spread Message	3.48	3.37	3.37	3.27

* $p < .05$

** $p < .01$

Comparing the difference between Authenticity and others

The ANOVA and LSD posthoc test indicated that the Authenticity summary variable ($M = 4.76$) was significantly greater than both the Clout variable ($M = 4.29$, $p < .05$) and the Negative Emotional Tone variable alone ($M = 4.12$, $p < .01$) while not significantly different from the Analytical variable ($M = 4.50$, $p = .15$). Significant differences were not found among

the treatments on justification or likelihood to spread the message. ANOVAs were also run on additional variables, including the credibility of the source, skepticism, and negative word of mouth intentions. However, the mean scores across the treatments were not significantly different. Therefore, messages that employ a negative message with authentic tone are the most believable, whereas messages that only convey a negative emotional tone alone are seen as the least believable. These treatments are used in subsequent studies to create baseline treatments.

Study 2 examines the role of the Negative Cascade in enhancing the believability of the message and the likelihood of inoculating others through information sharing. Specifically, the negative information is paired with a Negative Cascade or with No Cascade. This study provides insights into the role of an availability cascade as the mechanism driving group inoculation. Additionally, the function of a positive community response is used to determine the importance of a Positive Cascade in influencing the choice not to share the information. This component provides evidence that availability cascades can also work for the good of an organization by developing positive herd inoculation. Study 2a utilizes a 2x2 model (Negative Tone /Authentic) (No Cascade/Negative Cascade), and Study 2b uses a 2x2 model (Negative Tone / Authentic) (Positive Cascade /No Cascade) to measure these effects. The cascade effects and hypotheses are tested using PROCESS Models 6 and 83.

Study 2

Pretest

To ensure the reliability and validity of the adapted Likert-type and semantic differential scales, the researcher conducted a pretest with a sample of 303 respondents on Amazon's MTurk. Respondents with a 95% work acceptance rate were considered for participation, were compensated to reduce misrepresentation, and a working panel of respondents approved from

previous studies, and pretests were used (Hulland and Miller, 2018; Sharpe Wessling, Huber, and Netzer, 2017).

The total pretest sample consisted of 59% male respondents, with the majority of respondents falling below 40 years of age (73.8%). Additionally, the respondents indicated that they often or exclusively purchase online (75.9%), use social media at least daily (79.4%), and participate actively in online discussions (45.4%). Acceptable reliability ($\alpha > 0.70$) (Netemeyer, Bearden, and Sharma, 2003; Peterson, 1994) was found for all measures. Model fit statistics indicated that the measurement model had a good model fit and was stable enough to create composite variables to pretest the hypotheses. The hypotheses were pretested using PROCESS Models 6 and 83. The results for each hypothesis were significant and supported the contribution of the research. To ensure that the findings did not capitalize on the chance of using one dataset, additional samples were retrieved and used to confirm the findings.

Procedure

The study implements an experimental design research methodology to test the hypothesized relationships. A total of 308 participants took the survey through an online U.S. consumer panel purchased from Qualtrics; 52% between the ages of 21-40, and 51.6% female). 67.5% of the respondents indicated that they purchase either frequently or exclusively online, use social media at least daily (73.4%), and participate actively in online discussions (42.6%). (See Table 3.5). Participants were given a brief overview of the study and asked for their consent. Once participants agreed, they were given instructions and randomly shown one of four manipulated scenarios.

Table 3.5 Study 2 Sample Description

Category	Total Collection (308)
Age	
18-20	21 (7%)
21-30	69 (22%)
31-40	69 (22%)
41-50	56 (18%)
51-60	46 (15%)
61 and over	47 (15%)
Gender	
Male	149 (48.4%)
Female	159(51.6%)
Social Media Usage	
Never	6 (2%)
Monthly	8 (3%)
Weekly	37 (12%)
Daily	191 (62%)
Hourly	66 (21%)
Social Media Discussion Activity	
Not very active	39 (13%)
Not active	48 (16%)
Neither active nor inactive	60 (19%)
Active	124 (40%)
Very active	37 (12%)
Online Shopping Frequency	
Never	1 (0%)
Rarely	15 (5%)
Sometimes	82 (27%)
Often	165 (54%)
Always	45 (15%)

Similar to the MANOVA study, each participant was told about the company, New Latitudes, and then shown the Negative-Tone or Authentic treatment from the original poster (See Appendix). Afterward, they were shown either No Cascade or a corresponding cascade treatment. The cascade treatments were developed by creating a faux Reddit forum that aims to discuss the original poster’s comment. Care was taken to ensure that both the Negative Cascade and Positive Cascade were seen as distinct. A control treatment featuring No Cascade was used to compare to the Negative or Positive Cascade. In Study 2a, participants were shown (Negative

Tone/ Authentic and No Cascade/ Negative Cascade). In Study 2b, participants were shown (Negative Tone/ Authentic and Positive Cascade/ No Cascade). All manipulations are available in APPENDIX B.

Participants were then asked to answer survey questions regarding believability of the message ($\alpha=.93$; Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000), justification in sharing the message ($\alpha=.97$, Inman and Zeelenberg, 2002), and the likelihood to spread the message ($\alpha=.95$, Lee and Ma, 2012). All constructs, items, and reliabilities are available in Table 3.6. A qualitative reading check question was asked about halfway through the survey to ensure participants were paying attention; four participants who failed the reading check question were excluded from the study, and their responses were not used. After participants answered all survey items, a thought listing was given to participants, which told them the purpose of the study and asked them if they had any other thoughts about the impact of online information sharing. All the thought listings were reviewed for any abnormalities, and no apparent problems were evident.

A confirmatory factor analysis was then performed using AMOS 26 to assess the unidimensionality, convergent validity, and discriminant validity of the latent constructs. The results of the analysis indicated an acceptable fit (Marsh, Hau, and Wen, 2004) of the model to the data ($\chi^2=118.79$, $df = 51$, $CFI = .99$, $TLI = .98$ $IFI = .99$, $RMSEA = .06$). **Error! Reference source not found.** shows a complete list of results from the CFA, along with composite reliability for each construct.

Table 3.6 CFA Results, t-values, and Cronbach's α

Scales	α	Standardized Factor Loadings	t-values
Believability – (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000)	0.93		
Highly believable		0.78	19.27
Absolutely true		0.84	22.21
Totally acceptable		0.76	18.05
Very credible		0.94	*
Completely trustworthy		0.94	30.49
The justifiability of the decision - (adapted from Inman and Zeelenberg, 2002)	0.97		
Strongly Justifiable		0.90	30.88
Very Logical		0.93	35.98
Acceptable		0.96	43.04
Reasonable		0.97	*
Likelihood to Spread the Message (Lee and Ma, 2012)	0.95		
I would share the message online		0.98	38.04
I would share the message when discussing the company		0.89	30.00
I would spread the information to others		0.98	*

Model fit statistics

Note: $\chi^2 = 118.79$, $df = 51$, $p < 0.01$, $\chi^2/df = 2.29$, incremental fit index (IFI) = 0.99, Tucker-Lewis Index (TLI) = 0.98, comparative fit index (CFI) = 0.99, root mean square error of approximation (RMSEA) = 0.06.

All factor loadings have a p-value < 0.01

* denotes a constrained relationship to 1.00 for identification

Discriminant validity among the constructs was assessed using Fornell and Larcker's (1981) criterion. All of the constructs achieved discriminant validity as the average variance extracted for each construct was higher than the squared correlations between any pairs of constructs (see Table 3.7 for the variance between constructs and AVEs). The potential for common method variance was addressed by introducing a latent common method factor (Podsakoff, MacKenzie, Lee, and Podsakoff, 2003) to ensure that the measures are adequately represented. The results indicated that the presence of the latent common method factor did not significantly change the model fit ($\Delta\chi^2/df1 < .01$). Therefore, common method bias does not appear to be a significant concern.

Table 3.7 Means, Standard Deviations, AVEs, and Shared Variance

Scale	Mean	SD	1	2	3
Believability	4.04	1.41	0.73		
Justifiability of the Decision	4.09	1.80	0.58	0.88	
Likelihood to Spread the Message	3.54	1.87	0.37	0.47	0.88

Note: Average Variance Extracted is listed in the diagonals of the table

Further, participants were asked a manipulation check question regarding the initial LIWC treatment and the presence of a cascade of information (rated from overwhelmingly negative to overwhelmingly positive, and cases with incorrect responses were not used in the analysis. The manipulation check ensured that the participants were aware of whether they were shown a forum containing a “Negative Cascade/ No Cascade” or “Positive Cascade/ No Cascade.” Respondents were asked questions using a semantic differential scale. The measures indicate whether the information provided was overwhelmingly negative (when compared to No Cascade) or positive when compared to No Cascade) about New Latitudes ($t_{\text{negative}} = -6.78, p < .001$) ($t_{\text{positive}} = 16.08, p < .001$), primary against or primarily for New Latitudes ($t_{\text{negative}} -5.46, p < .001$) ($t_{\text{positive}} = 16.65, p < .001$), and in agreement with the original poster or in disagreement ($t_{\text{negative}} = -6.33, p < .001$) ($t_{\text{positive}} = 12.57, p < .001$). Therefore, the manipulation was successful.

Results

The unstandardized coefficient for the path from the original message to believability in the Negative Cascade/ No Cascade group (0.45) is significant, $t = 2.67, p < 0.05$. Therefore, Hypothesis 1 is supported. Additionally, the results are consistent in the Positive Cascade/No Cascade group (.54, $t = 2.74, p < .01$), also confirming Hypothesis 1. The use of an authentic message conveying a negative tone, rather than a message conveying only a negative tone, significantly influences the believability of the message content. Therefore, those attempting to

sway others can more effectively do so when masking their true intentions with implied sincerity.

Next, the path from believability to justification was tested. The unstandardized coefficient for the path from believability to justification in the Negative Cascade/ No Cascade collection is significant, 1.03, $t = 18.67$, $p < 0.01$. Thus, Hypothesis 2 is supported. Additionally, the relationship between believability and justification Negative Cascade/ No Cascade collection is also significant, (.94, $t = 15.50$, $p < .01$). Thus, Hypothesis 2 is supported again in the Positive Cascade condition. Consumers who see a message as believable are more likely to view the message content as justifiable in sharing.

Additionally, the level of justification significantly influences the likelihood to spread the message with others in both the Negative Cascade/ No Cascade and Positive Cascade/ No Cascade collections). The unstandardized coefficient (0.65) is significant, $t = 7.63$, $p < 0.01$, supporting Hypothesis 3. In addition, the Positive Cascade/ No Cascade collection, the unstandardized coefficient (.51, $t = 6.25$, $p < .01$) also supports Hypothesis 3. Messages that allow recipients to feel justified in sharing are more likely to be shared

Table 3.8 Studies 2a and 2b Results

	Hypothesized Relationship	Treatment	Unstandardized Estimates	t-Values	Direct Effect	Indirect Effect	CI: Low	CI: High	Hypothesis Supported?
	H1: Message Type => Believable	Negative Cascade	0.45	2.47*					Yes
		Positive Cascade	0.54	2.74**					
	H2: Believable => Justifiable	Negative Cascade	1.03	18.67**					Yes
		Positive Cascade	.94	15.50**					
	H3: Justifiable => Likelihood to Spread	Negative Cascade	0.65	7.63**					Yes
		Positive Cascade	0.51	6.25**					
Mediation	H4: Message Type => Believable => Vigilantism => Likelihood to Spread	Negative Cascade			0.03	0.30	0.06	0.57	Yes
		Positive Cascade			0.03	0.26	0.07	0.51	
Index of Moderated Mediation	H5a: Message Type X Negative Cascade * Believable* Justify => Share Self-created Misleading Information	Negative Cascade	0.80	2.20*		0.54	0.05	1.09	Yes
Index of Moderated Mediation	H5b: : Message Type X Positive Cascade * Believable* Justify => Share Self-created Misleading Information	Positive Cascade	-0.70	-2.07*		-0.38	-0.85	-0.01	Yes

**p<.01

*p<.05

Mediation

The mediating effects for Studies 1a and 1b are tested using PROCESS Model 6. The moderated mediation effects are then tested by introducing the moderator and using Model 83. The bootstrap method with 5,000 samples and a 95% confidence interval was used to calculate the mediating relationships (Hayes, 2018). The results of the mediation bootstrap analysis conducted using PROCESS are found in Table 3.8.

Message Type => Believable => Justifiable=> Spreading Message

Negative Cascade/ No Cascade Collection

The a-path (0.45, $t = 2.47$, $p < 0.05$) from the message type to believability, d-path from believability to justification (1.03, $t = 18.67$, $p < 0.01$), and b-path (0.65, $t = 7.63$, $p < 0.01$) from justification to the likelihood to spread the message were all significant. The indirect effect of the message type on the likelihood of sharing information through the mediators, believability, and justification, was also significant ($a \times d \times b = 0.30$, $p < 0.01$, CI Lower = 0.06, CI Upper = 0.57) (Zhao et al., 2010; Hayes, 2018). Since the confidence interval does not pass through zero, the indirect effect is determined to be statistically significant (Hayes, 2018). Additionally, since the direct effect of the message type on justification ($c_1 = 0.08$, $t = 0.66$, $p = 0.51$), believability on the likelihood to spread the message ($c_2 = 0.18$, $t = 1.67$, $p = 0.10$), and message type on the likelihood of spreading the message are all non-significant ($c = 0.03$, $t = 0.16$, $p = 0.87$), full, indirect-only serial mediation is found. Therefore, Hypothesis 4 is supported; believability and justification drive the relationship between the message type and the likelihood of spreading the message through a full serially mediated relationship.

Positive Cascade/ No Cascade Collection

The a-path (0.54, $t = 2.74$, $p < 0.01$) from the message type to believability, d-path from believability to justification (.94, $t = 15.50$, $p < 0.01$), and b-path (0.51, $t = 6.25$, $p < 0.01$) from justification to the likelihood to spread the message were all significant. The indirect effect of the message type on the likelihood of sharing information through the mediators, believability, and justification, was also significant ($a \times d \times b = 0.26$, $p < 0.01$, CI Lower = 0.07, CI Upper = 0.51) (Zhao et al., 2010; Hayes, 2018). Since the confidence interval does not pass through zero, the indirect effect is determined to be statistically significant (Hayes, 2018). Additionally, since the direct

effect of the message type on justification ($c_1 = -0.18$, $t = -1.06$, $p = .28$), believability on the likelihood to spread the message ($c_2 = 0.1$, $t = 1.60$, $p = 0.11$), and message type on the likelihood of spreading the message are all non-significant ($c = 0.034$, $t = 1.71$, $p = 0.09$), full, indirect-only serial mediation is found. Therefore, Hypothesis 4 is supported; believability and justification drive the relationship between the message type and the likelihood of spreading the message through a full serially mediated relationship.

Moderated Mediation

Negative Cascade

The 5,000 bootstraps with a 95% confidence interval used in the mediation analysis were also employed when calculating the moderated mediation relationships using Model 83 (Hayes, 2018; Zhao et al., 2010). The mediated relationship that was found between the message type on the likelihood of spreading the message through believability and justification was reviewed once again by including a Negative Cascade of confirmatory information. The interaction effect for the message type \times Negative Cascade is significant (0.80 , $t = 2.20$, $p < 0.05$). The interaction effect suggests that those exposed to an authentic message with a confirmatory Negative Cascade find the original post to be more believable. However, this effect was not found when the original poster did not confer authenticity in the original post. Additionally, the index of moderated mediation (0.54 , CI Lower = 0.05 , CI Upper 1.09) indicates that the presence of a Negative Cascade does moderate the full indirect effect on the likelihood of spreading the message. Therefore, Hypothesis 5a is supported. Consumers exposed to a negative message masked in authenticity are more likely to spread it if exposed to a Negative Cascade of confirmation information.

Positive Cascade

The mediated relationship that was found between the message type on the likelihood of spreading the message through believability and justification was also reviewed by including a Positive Cascade of disconfirmatory information. The interaction effect for the message type X Positive Cascade is significant ($-0.70, t = -2.07, p < 0.05$). The interaction effect suggests that an original post only conveying negative paired with a disconfirmatory Positive Cascade reduced the believability of the original poster. However, this effect was not present when participants were exposed to authentic message paired with a Positive Cascade. Additionally, the index of moderated mediation ($-0.38, CI Lower = -0.85, CI Upper -.01$) indicates that the presence of a Positive Cascade does also moderate the full indirect effect on the likelihood of spreading the message; therefore, Hypothesis 5b is supported. Consumers exposed to a negative message masked in authenticity are less likely to spread it if exposed to a Positive Cascade of disconfirmation information.

The full model with results can be found in Figure 3.3.

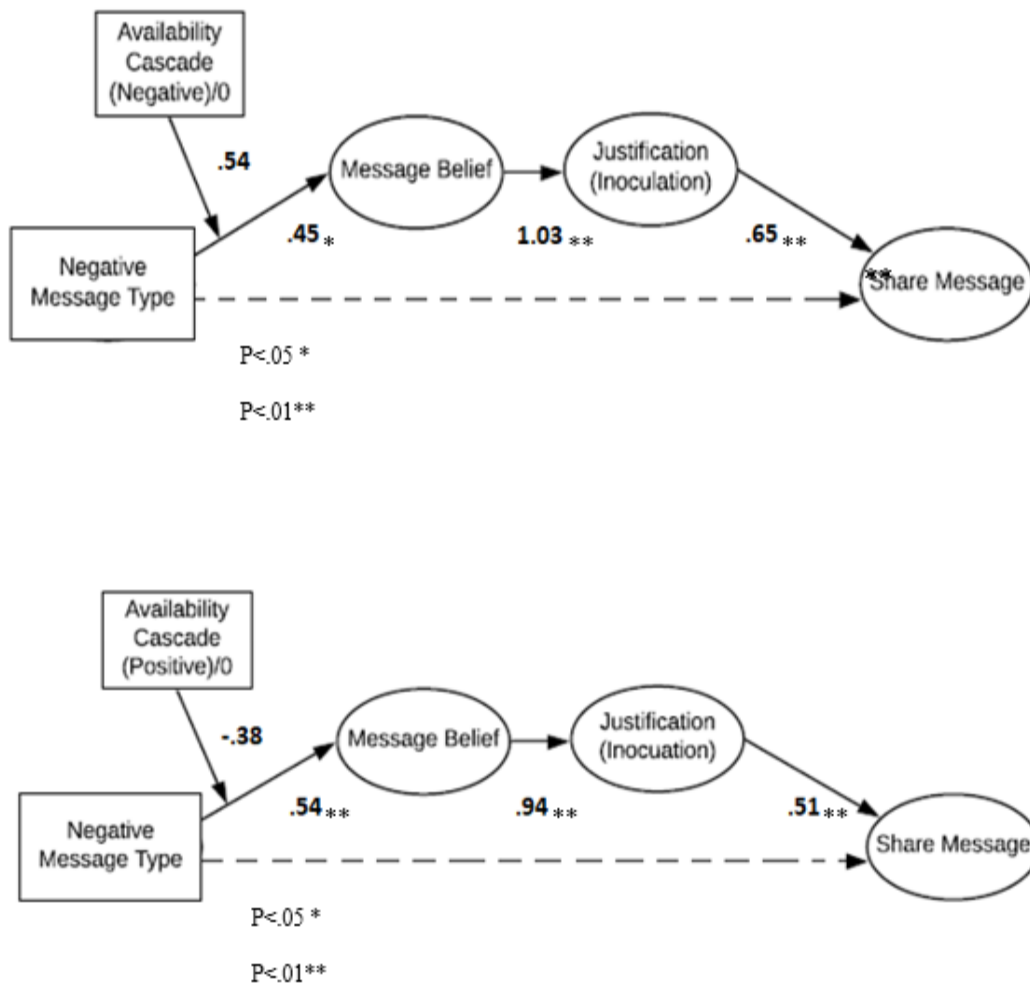


Figure 3.3 Study 2a and 2b Model Results

Discussion

The results from Studies 1a and 1b support H1-H5. The results further the findings of the MANOVA in showing that more authentic-seeming negative messages, rather than those exhibiting negative emotional tone alone, are more believable. When the message is more believable, it increases the likelihood of message sharing being perceived as justifiable. This higher level of justification strengthens the likelihood that a message is then shared with others,

leading to greater herd inoculation. Studies 1a and 1b add to the previous findings by showing an interaction between the original authentic message and cascades of negative (confirmation) or positive (disconfirmation) information. Negative authentic messages overall are seen as more believable than negative emotional tone messages; When the original message is paired with a Negative Cascade of information that shares similar sentiments, those viewing the forum are more likely to share the original message with others and drive negative herd inoculation. However, when a Positive Cascade that challenges the content of the original message accompanies the original message, others are much less likely to share the message. However, the Positive Cascade is most effective when the original post is less authentic. Therefore, a malicious post with the guise of authenticity may be shared online. These findings suggest that a Negative Cascade of information and negative herd inoculation can be attenuated when enough participants speak up to challenge a negative message posted against a brand. However, increasing the necessary participation to offset the misinformation can become increasingly challenging since consumers may resort to a truth default heuristic when the original poster conveys authenticity (Levine, 2016)

By understanding the role of availability cascades in driving the intentions to share misleading information, a better understanding of the cognitive processes and effort required to accept and share information when developing herd inoculation is uncovered. However, additional understandings of the roles of brand communities and the ways that brands can organize a response in a crisis are needed. Since the reaction of the firm will likely require more coordination than those sharing false beliefs, another study is needed to understand the necessity and challenges of coordinating an effective community response.

Study 3 will examine the presence of the bystander effect in reducing a consumer's likelihood to share information that defends a brand. Therefore, this study highlights the challenges faced by companies in overcoming negative consumer-driven inoculation. The respondents are exposed to an authentic/negative message, then the Negative/Positive Cascade. They are then asked if they would create information to post and asked if they would post a negative message in the forum or if they would be likely to post a positive message in the forum. The effects will then be compared to determine whether consumers are more likely to create negative or positive information. Study 3 uses structural equation modeling to allow for a multi-method validation of the model and the associated effects.

Consumer Bystanders

Consumer communities and brand communities can advocate for a brand offering to protect loved brands by actively disputing negative information and sharing positive impressions linked to the brand (Kravetz, 2007; Wallace, Buil, and Chernatony, 2014). These consumers may act as active advocates for other consumers by diligently disputing negative information that appears to be fraudulent (Wakefield and Inman, 1993; Larson and Denton, 2014). When these consumers pool informational resources together, shared insights can act to enhance or improve the overall brand (Schau, Muñiz, and Arnould, 2009; Wirtz et al., 2013).

Additionally, brand and consumer advocacy may distinguish offerings from those of others and continually build toward an active collaboration between the brand and its constituents (Fournier and Lee, 2009; Kravetz, 2007). By accessing the shared knowledge of a vibrant community, individual consumers are exposed to an abundance of positive information from brand advocates who act to protect the slandered brand. Even with an active community, though, the majority of members may never speak up or repost.

The bystander effect explains the reason that people may choose not to act in situations where others –even loved brands – are the target of negative word of mouth (Fischer et al., 2011). The term “bystander effect” (also called diffusion of responsibility) originates with the murder of Kitty Genovese in the 1960s’ in which numerous neighbors witnessed the crime taking place but did not call the police. The bystander effect is the tendency for even well-intentioned people to assume that someone else will surely take the initiative and act (Fischer et al., 2011). However, this is not always the case. Although the bystander effect can and does occur in emergency situations, this effect does not always rely on an emergency or a highly threatening situation (Levy et al., 1972). Bystander roles have been studied in social settings (Levine and Crowther, 2008), work-setting harassment (Johnson and Smith, 2017), bullying (Machackova, Dedkova, and Mezulanikova, 2015; Salmivalli, 2019); and bystander roles can be implicit in both hypothetical settings/ or imagined settings (Garcia, Weaver, Moskowitz, and Darley, 2002). In the context of malicious information sharing, the bystander effect describes consumers’ cognitive inclination to step back and let others defend the brand. Brand communities can be a significant asset for a marketer, but ensuring that the members will mobilize to offset harmful information sharing can be a daunting process.

Employing additional cognitive resources requires much more effort. Because of this effort requirement, predispositions and preferences that occur through automatic responses can then bias the accuracy of the analytical cognitive judgments and responses (Kahneman, 2011). For instance, online community members may assume that the negative information posted about a brand is not worth arguing against since, if it were not truthful, then other posters would surely challenge it. This lack of response is called the *implicit* bystander effect in that individuals are

implied as bystanders when they choose not to act in a given situation when other parties are present (Garcia et al., 2002). A resulting lackluster response could suggest that the bystander effect presents itself when consumers are exposed to messages that confirm the original poster's point of view (Fischer et al., 2011) and observing consumers fail to devote cognitive resources to assessing the situation to counter the information. Therefore, automatic responses are employed to minimize the use of working memory (Evans and Stanovich, 2013).

Believability of the Message

Since the original message acts as the initiating message for inoculation, the corresponding cascade of information serves to implement the delay phase of inoculation. This delay, required for successful persuasion, suggests that the original poster stop information sharing before malicious intentions become overtly obvious (Banas and Rains, 2010). Therefore, other posters dictate whether the viewpoint is spread.

When the source of the message is congruent with the tone in community response, then a message can be deemed believable. In this situation, the inoculation process flows through the more automated aspects of decision making since more considerable cognitive effort is not required. However, the original message may also be contradicted by the sentiments of a community. When this contradiction occurs, the tone of the forum can be incongruent with the original message.

Thus, the inoculation then requires both the initial automatic response with additional cognitive resources as are necessary for decision-making. When this occurs, the believability of the original message is diluted, and credibility is more difficult to decipher. Therefore, consumer bystanders are more likely to believe the original message in the presence of a confirmation rather than a disconfirmation cascade of information.

Hypothesis 6: The message from the original poster will positively influence a consumer's belief of the negative message in the presence of a Negative Cascade of information more so than in the presence of a Positive Cascade of information

Justification

The discussion itself becomes increasingly important in determining the perceived legitimacy of a concern (Kozinets, 2015; Kozinets, 2016). When the initial "gut" reaction is to believe a statement, and this reaction is confirmed by the discussions of a larger group, then the consumer bystander may agree with the sentiments shared. The individual may then determine that the overall tone is justified since the brand is painted in a negative light by all parties. Thus, the more automatic components may bias the justification in the tone. However, when the tone disconfirms the original poster, the tone itself may also be seen as justifiable since individuals may defend a brand. However, this justification in tone may not be as strong since the conflicting information may make decision-making more difficult.

Consumers tend to place greater importance in negative reviews (Nazlan, Tanford, and Montgomery, 2018). When viewpoints are consistent, less effort is required in decision-making (Guzmán, Paswan, and Van Steenburg, 2015). For instance, a negative forum that confirms a negative poster makes sense to others since it is congruent and consistent with expectations. Those that experience conflicting messages will likely feel more uncertain about actions that should be taken (Bee, and Madrigal, 2013) and incongruences make the motive for the forum response less clear. Since greater levels of justification aid in action, those who receive conflicting information will be less likely to buy into the tone (overarching attitudes) of the group discussion. When the content of information conforms to an availability heuristic, justifying the tone of the forum can be more easily explained to oneself and others. Thus, when

a message is deemed believable, and confirmed by the negative tone of a larger group of consumers, it is therefore considered as more justifiable. Therefore:

Hypothesis 7: The believability of the message will positively influence a consumer's justification in the forum tone in the presence of a Negative Cascade of confirmatory information more so than in the presence of a Positive Cascade of disconfirmatory information.

Posting a Positive Message

Once the consumer bystander perceives the forum tone to be justifiable, the individual can then exhibit propensities to get involved. When the negative message is later supported by corresponding negative discussions, then the consumer bystander will choose not to post anything positive to defend a brand since the automatic decision-making and heuristics from availability support the idea that the brand is guilty of some transgression (Levine, 2014). However, those exposed to a Positive Cascade that disconfirms the original message exhibit greater cognitive resources from comparison and will be more likely to post a positive message in the forum. However, this effect is expected to be weaker than those exposed to a Negative Cascade. When this occurs, the consumer bystander will likely not defend a brand that is being attacked, and s/he will be less likely to act to praise the brand than those condemning the brand. Therefore:

Hypothesis 8: The justification of the forum tone will negatively influence a consumer's likelihood to post a positive message in the presence of a Negative Cascade of confirmatory information more so than in the presence of a Positive Cascade of disconfirmatory information

Posting a Negative Message

Since the consumer bystander perceives the forum tone to be justifiable, the individual can also contribute to the negativity of the discussion. Since the automatic decision-making and heuristics from availability suggest that the brand is guilty of some transgression, posting a negative message can be easily justified (Levine, 2014). When the negative message is then confirmed by corresponding negative discussions, then the “would be” consumer bystander will contribute to the growth of the Negative Cascade and strengthen herd inoculation. However, those exposed to will be unlikely to post anything negative since doing so could be challenging to justify, and the evidence that drives the automatic and analytical resources used in decision-making is not found. When this occurs, the consumer bystander will likely not defend a brand that is being attacked nor contribute to the negative information. Therefore:

Hypothesis 9: The justification of the forum tone will positively influence a consumer’s likelihood to post a negative message in the presence of a Negative Cascade of confirmatory information more so than in the presence of a Positive Cascade of disconfirmatory information

The Bystander Effect

Since the herd inoculation requires multiple parties, it becomes increasingly important to address inoculation’s influence in future information sharing. Since an availability cascade can reinforce the legitimacy of a message online by suggesting greater trustworthiness of the message (Kane and Webster, 2012), it then also contributes to future information sharing and the creation of messages that confirm the message. Additionally, the presence of a Negative Cascade of confirmation information allows a consumer to resort to more heuristic or automatic decision-making since the information can reinforce the original message (Kahneman, 2011; Kane and

Webster, 2012). Thus, confirmatory negative information is expected to influence the inoculation of the consumer bystander since the automatic responses dictate the evaluation of information credibility through availability and enhance the herd inoculation process. When the consumer bystander deems the information of the original poster and community to be congruent, then s/he will be more likely to get actively involved in the forum discussion.

When consumers view messages that are inconsistent, additional cognitive resources may be necessary to discredit the original poster or the community. These consumers may then immediately assume the original message was false, but they may also not get involved to further embolden the discussion by writing a negative statement. They, however, will be likely to post a positive comment within the forum but with less enthusiasm.

When the cascade of information confirms the original post, the consumer bystander will be less likely to post positive information and more likely to post negative information in the forum. However, when the cascade disconfirms the original post, the additional cognitive resources required weakens the likelihood to post a positive post. The presence of a Negative Cascade is expected to influence a mediated relationship between the original poster's message and the likelihood of spreading the message belief and justification in message sharing.

Therefore:

Hypothesis 10a: A serially mediated relationship between the original message and the likelihood to post a *positive* message, through the belief in the message and perceived justification of the forum tone, will be stronger in the presence of a Negative Cascade than the presence of a Positive Cascade.

Hypothesis 10b: A serially mediated relationship between the original message and the likelihood to post a *negative* message, through the belief in the message and

perceived justification of the forum tone, will be stronger in the presence of a negative cascade than in the presence of a Positive Cascade.

See the full conceptual models for Study 3 in Figure 3.4 below.

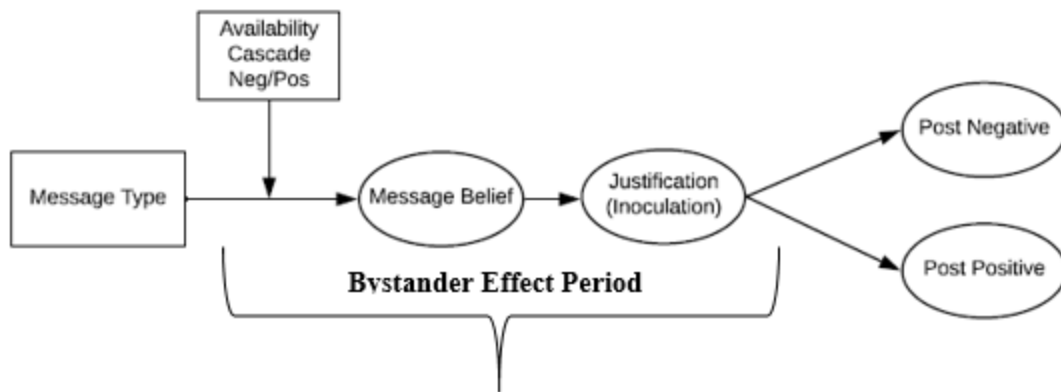


Figure 3.4 Study 3 Conceptual Model

Study 3: The Bystander Effect

The constructs used in the previous studies were used to test the bystander effect. The scale items presented both reliability and validity in previous studies and are deemed appropriate for use in the current model. A pretest was first conducted to determine the validity and reliability of the items as applied to the present study, and no issues were found with reliability, validity, or unidimensionality.

Sample and Procedure

The study employs an experimental design research methodology to test the hypothesized relationships. The researcher recruited 574 respondents for the confirmatory study. To ensure quality online respondents, the study included various attention checks, respondent screening

policies, and a requirement of written text. Also, only U.S. respondents with a 95% work acceptance rate were considered for participation, a nominal wage was paid to reduce misrepresentation, and a working panel of respondents approved from previous studies, and pretests was used (Hulland and Miller, 2018; Sharpe Wessling, Huber, and Netzer, 2017). Respondents who missed attention measures were not allowed to continue the study. Participants for the Negative Cascade treatment (291 participants) and for the Positive Cascade treatment (283 participants) were recruited from M-Turk who had purchased from companies online, and who frequently use and participate in social media (Peterson and Merunka, 2014). (See Table 3.9)

Table 3.9 Study 3 Sample Description

Category	Total Collection (574)
Age	
18-20	4 (0%)
21-30	148 (26%)
31-40	194 (34%)
41-50	101 (18%)
51-60	74 (13%)
61 and over	53 (9%)
Gender	
Male	270 (47%)
Female	304 (53%)
Social Media Usage	
Never	19 (3%)
Monthly	34 (6%)
Weekly	85 (15%)
Daily	365 (64%)
Hourly	71 (12%)
Social Media Discussion Activity	
Not very active	74 (13%)
Not active	156 (27%)
Neither active nor inactive	89 (16%)
Active	219 (38%)
Very active	36 (6%)

New Table 3.9 (continued)

Online Shopping Frequency	
Never	1 (0%)
Rarely	22 (4%)
Sometimes	153 (27%)
Often	337 (59%)
Always	61 (11%)

Participants were exposed to information about an online retailer called New Latitudes that is then accused of deleting negative reviews. The participants are then exposed to one of two treatments (a negative confirmation cascade or a positive disconfirmation cascade of information). After exposure to the condition, respondents completed an online questionnaire (via Qualtrics). The scale items from Study 2a and 2b, related to the believability of the message (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000) and justification in sharing the message (Inman and Zeelenberg, 2002), are used in this study. The likelihood to post positive or negative information (Brüggen, Foubert, and Gremler, 2011; Reichheld, 2003) and the likelihood to participate in the discussion (Grau and Folse, 2007) scales were adapted from prior research. The sample consisted of 53% male respondents with a mean age of 40 years. 69.6% of the respondents indicated that they purchase either frequently or exclusively online, use social media at least daily (76.2%), and participate actively in online discussions (44.8%).

A manipulation check ensured that the participants were aware of whether they were shown a forum containing a “Negative Cascade” or “Positive Cascade.” Respondents were asked questions using a semantic differential scale to indicate whether the information provided about New Latitudes was overwhelmingly negative or positive ($t = -38.21, p < .001$), primary against or primarily for New Latitudes ($t = -36.92, p < .001$), and in agreement with the original poster or in

disagreement ($t = -38.64, p < .001$) . Based on the responses, the manipulation check was successful.

A preliminary bystander effect was also found for those exposed to the Positive Cascade versus the Negative Cascade. Those exposed to the negative information were more likely to participate in the forum $t = 2.07, p < .05$. This finding suggests that participants should be more likely to get involved in a discussion when negative information is being shared rather than positive.

Confirmatory Factor Analysis

Tests using AMOS 26 showed that the composite reliability of all constructs exceeded 0.80, indicating evidence of construct reliability (Garver and Mentzer, 1999). The CFA fit statistics for online consumers show a satisfactory model fit. The χ^2 of 384.74 with 142 degrees of freedom ($p < 0.01$) provides a χ^2/df ratio of 2.70. The IFI = 0.98, TLI = 0.98, CFI = 0.98 and RMSEA = 0.05, each providing additional evidence of acceptable model fit (Hu and Bentler, 1999). Also, the t-values (found in **Error! Reference source not found.**) are all significant ($p < 0.01$). The AVEs for each construct exceed 0.50, providing evidence of convergent validity (See **Error! Reference source not found.**). The AVEs for each construct exceeds the level of shared variance between constructs, providing evidence of discriminant validity (Fornell and Larcker, 1981). The potential for common method variance was addressed by introducing a latent common method factor (Podsakoff, MacKenzie, Lee, and Podsakoff, 2003) to ensure that the measures are adequately represented. The results indicate that the presence of the latent common method factor did not significantly change the model fit ($\Delta\chi^2/df1 < .01$). Therefore, common method bias does not appear to be a major concern.

Table 3.10 CFA Results, t-values, and composite reliability

Scales	c.r	Standardized Factor Loadings	t-values
Believability – (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000)	0.94		
Highly believable		0.80	26.89
Absolutely true		0.87	32.36
Totally acceptable		0.84	29.77
Very credible		0.92	*
Completely trustworthy		0.92	37.47
The justifiability of forum tone - (adapted from Inman and Zeelenberg, 2002)	0.94		
Strongly Justifiable		0.86	33.19
Very Logical		0.87	34.27
Acceptable		0.92	40.74
Reasonable		0.94	*
Post Positive Information - (adapted from Brügger , Foubert, and Gremler, 2011; Reichheld, 2003)	0.94		
I would say positive things about New Latitudes in the forum		0.91	35.55
I would write a message of approval for New Latitudes to others in the forum		0.93	*
I would encourage other participants in the forum to support New Latitudes		0.90	35.07
Post Negative Information - (adapted from Brügger, Foubert, and Gremler, 2011; Reichheld, 2003)	0.95		
I would say negative things about New Latitudes in the forum		0.92	44.51
I would write a message rejecting New Latitudes in the forum		0.97	*
I would discourage other participants in the forum from supporting New Latitudes		0.89	38.78
Likelihood to participate in the Discussion - (adapted from Grau and Folse, 2007)			
I would be willing to participate in this forum	0.96	0.93	39.78
I would consider involving myself in the forum discussion		0.95	45.32
It is likely I would contribute my thoughts in the forum		0.93	*
I think getting involved online in the forum is a good idea		0.89	35.36

Model fit statistics

Note: $\chi^2 = 384.74$, $df = 142$, $p < 0.01$, $\chi^2/df = 2.71$, incremental fit index (IFI) = 0.98, Tucker-Lewis Index (TLI) = 0.98, comparative fit index (CFI) = 0.98, root mean square error of approximation (RMSEA) = 0.5.

All factor loadings have a p-value < 0.01

* denotes a constrained relationship to 1.00 for identification

Table 3.11 Means, Standard Deviations, AVEs, and Shared Variance

	Mean	SD	1	2	3	4	5
Believability	4.48	1.30	0.76				
Justifiability of Forum Tone	4.98	1.37	0.21	0.81			
Post Positive Information	3.33	1.43	0.08	0.01	0.83		
Post Negative Information	3.37	1.51	0.20	0.07	0.00	0.86	
Likelihood to Participate	3.81	1.75	0.05	0.13	0.12	0.23	0.85

Note: Average Variance Extracted is listed in the diagonals of the table

Invariance

After the CFA is completed for each group, configural invariance is established by testing the unconstrained model. Configural invariance ensures that the factor structures are equivalent between the groups in the unconstrained model and must be established before proceeding to further analysis. $\chi^2 = 3117.86$ and $df = 609.13$, χ^2/df ratio of 2.15 provides evidence of a good model fit. The IFI = .97, TLI = .97, CFI = .97 and RMSEA = .05 provide further evidence of configural invariance across the groups in the unconstrained model.

Next, metric invariance is determined by constraining factor loadings to be equal across groups. Metric invariance ensures that the meaning of the construct is equivalent across the groups. The measurement weights are constrained across the two treatments to provide evidence of non-significance in group differences. The $\chi^2 = 15.30$, $df = 14$, $p = .26$, and TLI = .02 provide evidence that the difference between the two groups is non-significant. Evidence of full metric invariance is found.

Structural Equation Modeling

Structural equation modeling provides flexibility in model creation, can reduce bias in measurement, and provides model fit statistics while calculating the moderated mediation effects when proper measurement and structural models are specified (Hayes, Montoya, and Rockwood, 2017). After the assessment of the measurement model, the full model was analyzed across both

treatments using AMOS 26 (See Table 12). The results of the structural model provide evidence of an acceptable model fit ($\chi^2 = 433.62$, $df = 101$, $\chi^2/df = 4.29$, $p < 0.01$, IFI = 0.96, TLI = 0.95, CFI = 0.96 and RMSEA = 0.07) (Garver and Mentzer, 1999; Marsh et al., 2004; Lomax and Schumacker, 2004). Therefore, the structural model, mediation, and moderated mediation analyses can be confidently conducted. The standardized coefficients are compared across each group (negative or positive) to compare the strength of the relationships.

Results

Message Type and Believability

The standardized coefficient for the path from the original message to believability in the Negative Cascade treatment (0.17) is significant, $t = 2.78$, $p < 0.01$. Additionally, the results are not consistent in the Positive Cascade treatment (.07, $t = 1.20$, $p = .23$). Therefore, Hypothesis 6 is supported. The use of an authentic message conveying a negative tone rather than a message conveying only a negative tone significantly influences the believability of the message content when paired with a confirmatory Negative Cascade rather than a disconfirmatory Positive Cascade. Therefore, those attempting to sway others can more effectively do so when masking their true intentions with sincerity if others are posting in a similar tone.

Believability and Justification in Forum Tone

Next, the path from believability to justification was tested. The standardized coefficient for the path from believability to justification in the Negative Cascade collection is significant, .77, $t = 13.83$, $p < 0.01$. Additionally, the relationship between believability and justification in the Positive Cascade collection is also significant (.16, $t = 2.61$, $p < .01$). A χ^2 difference test was conducted to determine if the paths in the Negative Cascade and Positive Cascade treatments

were significantly different. The resulting χ^2 (46.62, $p < .001$) indicates that a significant difference is present, supporting Hypothesis 7. Consumers who see a message as believable are more likely to view the forum tone as justifiable in sharing. When a bystander is exposed to a negative message with a confirmatory Negative Cascade in a forum, they are much more likely to feel that the tone of the forum is justified rather than those exposed to a disconfirmatory Positive Cascade in a forum.

Justification in Forum Tone and Posting Positive Information

The level of justification in the forum tone becomes increasingly important in determining the likelihood to join in the discussion. This finding can significantly influence the likelihood to post positive information about a brand in the forum. The standardized coefficient for the Negative Cascade group (-0.25) is significant, $t = -4.06$, $p < 0.01$). This finding suggests that consumer bystanders are significantly less likely to post positive information about a brand in the presence of a cascade of negative information. In addition, the Positive Cascade the standardized coefficient (.20, $t = 3.16$, $p < .01$) is also significant. A χ^2 difference test was conducted to determine if the paths in the Negative Cascade and Positive Cascade treatments were significantly different. The resulting χ^2 (26.47, $p < .001$) indicates that a significant difference is present. Thus, Hypothesis 8 is supported. This finding indicates that consumers exposed to an initial negative message about a brand that is then disputed by a group of consumers are more likely to post positive posts about a brand in the forum. However, those exposed to an initial negative message a negative confirmatory cascade have no intention of posting a positive message since herd inoculation is beginning to take effect. Additionally, this effect is stronger in the presence of the Negative Cascade than the Positive Cascade. The actions of the consumers in posting a positive post are seemingly closely tied to the tone of the forum.

Justification in Forum Tone and Posting Negative Information

The justification in the forum tone is also essential to determine the likelihood to join in and post negative information. This can significantly influence the possibility to post negative information about a brand in the forum and reinforce, and even grow, the cascade. The standardized coefficient for the Negative Cascade group (0.46) is significant, $t = 7.75$, $p < 0.01$. This finding suggests that consumer bystanders are significantly more likely to post negative information about a brand in the presence of a cascade of negative information. The Positive Cascade standardized coefficient (.03, $t = .55$, $p = .58$) is non-significant.

A χ^2 difference test was conducted to determine if the paths in the Negative Cascade and Positive Cascade treatments were significantly different. The resulting χ^2 (31.96, $p < .001$) indicates that a significant difference is present and supports Hypothesis 9. This finding suggests that consumers exposed to an initial negative message about a brand that is then confirmed by a group of consumers display a greater propensity to participate in the forum . However, a consumer who is exposed to a Positive Cascade is unlikely to be influenced to post negative information and further the herd inoculation. The consumers who are exposed to Negative Cascades of information are more likely to get involved by posting negative information. (See results in Table 3.12 and 3.13).

Table 3.12 Study 3 Results

Hypothesized Relationship	Negative Cascade		Positive Cascade		Two Group Difference $\Delta\chi^2/1df$	Hypothesis Supported?
	Standardized Estimates	t-Values	Standardized Estimates	t-Values		
H6: Message Type => Believable	0.17	2.78*	0.07	1.20		Yes
H7: Believable => Justifiable	0.77	13.83**	0.16	2.61**	46.62 (p<.001)	Yes
H8: Justifiable => Post Positive	-0.25	-4.06**	0.20	3.16**	26.47 (p<.001)	Yes
H9: Justifiable => Post Negative	0.46	7.75**	.03	.55	31.96 (p<.001)	Yes

Table 3.13 Mediation Analysis

		Direct	Indirect	Low	High	Two Group Difference Δadc	Hypothesis Supported?
Negative Cascade	H10a: Message=>Justifiable=>Post Positive	-0.35*	-0.07**	-0.18	-0.02	-0.08**	Yes
Positive Cascade	H10a: Message=>Justifiable=>Post Positive	-0.21	0.001	-0.002	0.02		
Negative Cascade	H10b: Message=>Justifiable=>Post Negative	0.25	0.16**	0.05	0.31	.17**	Yes
Positive Cascade	H10b: Message=>Justifiable=>Post Negative	.14	0.006	-0.001	0.03		

$\chi^2 = 433.62$, $df = 101$, $\chi^2 / df = 4.29$, $p < 0.01$, incremental fit index (IFI) = 0.96, Tucker-Lewis Index (TLI) = 0.95, comparative fit index (CFI) = 0.96, root mean square error of approximation (RMSEA) = 0.07 **indicates $p < .01$, * indicates $p < .05$)

Mediation

The bootstrap method with 5,000 samples and a 95% confidence interval was used to calculate the mediating relationships (Hayes, 2018). The results of the mediation bootstrap analysis conducted using AMOS 26 are found in Table 3.12. The mediating effects in Study 3

are calculated by creating estimands to estimate the indirect effects of all paths using structural equation modeling. Unstandardized results are used to calculate indirect effects.

Message Type => Believable => Justifiable=> Post Positive

Negative Cascade

The significant a-path and c-path used above are used to calculate the mediating effects for the likelihood to share negative information in the Negative Cascade group. The b-path (-0.24, $t = 7.45$, $p < 0.01$) from justification to the likelihood to post a positive message was significant. The indirect effect of the message type on the probability of sharing information through the mediators, believability, and justification, was also significant ($axdxb = -.07$, $p < 0.01$, CI Lower = $-.18$, CI Upper = $-.02$) (Zhao et al., 2010; Hayes, 2018). Since the confidence interval does not pass through zero, the indirect effect is determined to be statistically significant (Hayes, 2018). The direct effect of the message type on justification ($c = -.35$, $t = -2.17$, $p < .05$) was also significant. Therefore, believability and justification reduce the relationship between the original message and the likelihood to post positive information in the forum. This finding suggests that both the negative group dynamic and the original message reduce the likelihood to post positive information about the brand.

The a-path (0.40, $t = 2.77$, $p < 0.01$) from the message type to believability, d-path from believability to justification (.78, $t = 13.83$, $p < 0.01$), and b-path (0.53, $t = 7.45$, $p < 0.01$) from justification to the likelihood to post a negative message were all significant. The indirect effect of the message type on the likelihood of sharing information through the mediators, believability, and justification, was also significant ($axdxb = 0.16$, $p < 0.01$, CI Lower = 0.05 , CI Upper = 0.31) (Zhao et al., 2010; Hayes, 2018). Since the confidence interval does not pass through zero, the indirect effect is determined to be statistically significant (Hayes, 2018). Additionally, since the

direct effect of the message type on justification ($c=0.25$, $t= 1.56$, $p=0.12$), full mediation is found. Therefore, believability and justification drive the relationship between the original message and the likelihood to post negative information in the forum through a full serially mediated relationship. This finding suggests that the negative group dynamic, rather than the original message, influences the likelihood to post additional negative information and further continue herd inoculation.

Positive Cascade

In the presence of the Positive Cascade condition, the a-path (0.18 , $t = 1.21$, $p=.23$) from the message type to believability is non-significant. However, d-path from believability to justification ($.18$, $t = 2.62$, $p < 0.01$), and b-path (0.19 , $t = 3.15$, $p < 0.01$) from justification to the likelihood of posting positive information are significant. The indirect effect of the message type on the likelihood of sharing information through the mediators, believability, and justification, was, however, non-significant ($axdxb = 0.001$, $p=0.37$, CI Lower $=-0.002$, CI Upper $=0.02$) (Zhao et al., 2010; Hayes, 2018). Additionally, since the direct effect of the message type on the likelihood of posting a positive message is also non-significant ($c=-.21$, $t= -1.30$, $p=0.20$), neither a mediated nor direct effect is found to the likelihood to share positive information. Therefore, those exposed to a Positive Cascade of information are neither more likely to post a positive message, nor are they likely to post a positive message after an initial negative message.

The a-path and c-path are also used to calculate the presence of a mediated effect in the Positive Cascade group. The b-path (0.03 , $t = .56$, $p =.58$) from justification to the likelihood to post a negative message was, however, non-significant. Additionally, the indirect effect, ($axdxb = 0.006$, $p=0.08$, CI Lower $=-0.001$, CI Upper $=0.03$) was non-significant. This finding, along with the negative direct effect of the negative message on the likelihood to post a negative

message ($c = .14$, $t = .84$, $p = .40$) suggests that individuals may choose not to get involved when the initial message is negative nor when the forum discussion is protecting a brand. Therefore, those exposed to a Positive Cascade of information are neither more likely to post a negative message, nor are they likely to post a negative message after an initial negative message is viewed. These findings suggest that the bystander effect can occur when information is either incongruent or the individual feels that there is not a need to get involved in an online discussion. This inaction suggests that the role of the group in driving herd inoculation will continue without active challenges

Mediation Comparison

Moderated Mediation

Estimands are created to calculate the mediated effect in both the negative and Positive Cascade groups and then compare the differences. These estimands conduct the bootstrap calculations while estimating the interaction effects created while allowing for multiple DVs (Bayl-Smith and Griffin, 2014). The 5,000 bootstraps with a 95% confidence interval used in the mediation analysis were also employed when calculating the moderated mediation relationships (Hayes, 2018; Zhao et al., 2010).

To further explore if differences were present, a moderated mediation test was conducted to see if significant differences were present across the groups. The results of the analysis found that across the Negative Cascade and Positive Cascade groups, the mediation results significantly differ in predicting the posting of positive information. The difference in the indirect effect of the Negative Cascade group minus the indirect effect of the Positive Cascade is significant ($a \times d - b = .08$, $p < .01$, CI Lower = -0.18 , CI Upper = -0.03). Therefore, Hypothesis 10a is supported. The negative and Positive Cascades are significant in determining whether positive information is

shared in a forum. Those exposed to a Negative Cascade of information are much less likely to post positive information.

Additionally, the mediation results significantly differ in predicting the posting of negative information. The difference in the indirect effect of the Negative Cascade group minus the indirect effect of the Positive Cascade is significant ($axdx_b = .17$, $p < .01$, CI Lower = 0.05, CI Upper = .31). Therefore, Hypothesis 10b is supported as well. Those exposed to the Negative Cascade are significantly more likely to believe the initial message, find the tone of the forum justifiable, and post negative information in the forum.

Ultimately, this test of mediation emphasizes the importance of understanding the negative effects of Negative Cascades as well as the challenges in gaining the participation of consumer bystanders in Positive Cascades. In both situations, brands must find ways to attenuate the negative effects while encouraging positive information sharing (see Figure 3.4).

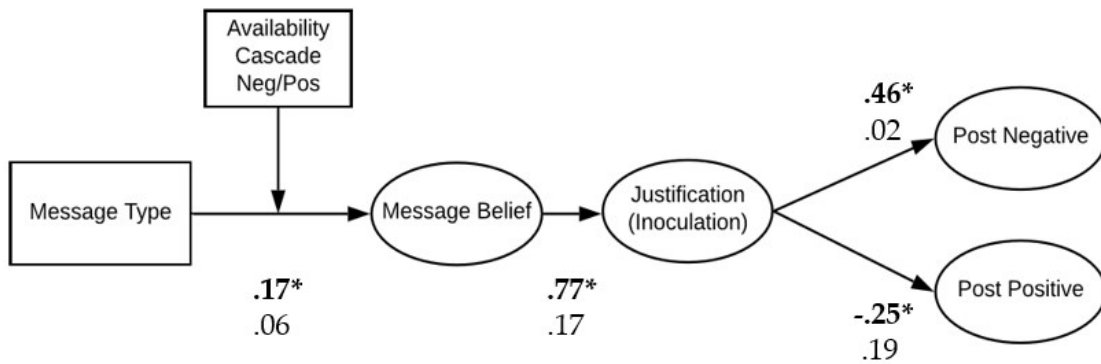


Figure 3.4 Study 3 Model Results

Discussion

The results of Study 3 further the findings of the Studies 1a and 1b by indicating that those exposed to the Negative Cascade were much more likely to get involved in the forum and further the role of herd inoculation. The original message is seen as more believable when others confirm it and when consumers can more easily justify the overall tone in the forum. This higher level of justification strengthens the likelihood that a negative message is created reduces the likelihood that a positive message is shared. Alternatively, a Positive Cascade reduces the believability and justification since the messages are in conflict. Although participants are more likely to post a positive message, the standardized effect is weaker than that of the Negative Cascade group. This finding provides evidence of a bystander effect in gaining when more cognitive resources are required to act. When the original message is paired with a Negative Cascade of information that shares similar sentiments, those viewing the forum are more likely to participate in the forum discussion and contribute to the herd inoculation. However, when the original message is paired with a Positive Cascade that challenges the content of the original message, bystanders are less likely to get involved.

By understanding the role of availability cascades in driving the intentions to share misleading information, a better understanding of the cognitive processes and effort required to accept and share information is developed. This understanding contributes to the knowledge of the bystander effect in protecting a brand from misleading information. If the action requires mental strain or effort, the consumer may be likely to deflect responsibility to others. Therefore, consumers online may resort to heuristics in that they do not feel the need to challenge the status quo of a forum discussion and choose to remain a bystander and allow the herd inoculation to go unchecked.

General Discussion and Implications

The preceding studies provide evidence that the original message type impacts whether consumer bystanders believe the message. Moreover, the message believability was the biggest differentiator between the Analytical, Clout, Authentic, and Negative Emotional Tone messages (more so than source credibility or skepticism). The thought listings that were included at the end of each survey in the individual studies provide additional insight. Participants stated that they felt like the original poster was trying to help when sharing a message with an authentic tone. This finding suggests that those attempting to sway others with malicious attempts can position themselves as a lion in sheep's clothing, masking their true intentions. Study 2a showed that consumers are more likely to share the original message and extend the herd inoculation when confirmed by a Negative Cascade. Study 2b showed that a Positive Cascade that disconfirms the original message reduces the likelihood to share and can ultimately stop the flow of negative information and halt the spread of herd inoculation and may even flip the direction of the inoculation. Study 3 also showed the importance of availability cascades in forum participation. Both the negative and Positive Cascade treatments provided evidence that individuals can develop intentions to participate. Thus, the main effects found in Studies 1a and 1b still hold. However, a Negative Cascade of information maintained the mediated relationship to increase the likelihood to post negative information and decrease the likelihood to post a positive message.

Since those exposed to the Negative Cascade stated that they would be more likely to participate in the forum, the cascade becomes increasingly crucial in determining the presence and strength of a bystander effect. Negative information is memorable (Schwarz, 2004) and reaffirms the heuristic shortcut developed after the original message (Meyers-Levy and

Maheswaran, 2004). When the Negative Cascade is then seen, it is much easier to join in and share the message and much more challenging to develop the necessary cognitive justification to fight the message. When the negative message belief is broken by contrary information, it requires a little more cognitive resources and the snap judgments to both participate in either a positive or negative way. However, the inaction of bystanders makes it much more difficult for brands to recover.

These findings have important and clear implications for managers as they can use the findings of these studies to focus on tracking mechanisms, causing misinformation sharing and responding to potential firestorms. Community participants, the herd, clearly determine whether or not a negative message picks up traction and spreads. Brands should diligently monitor online communities and encourage participation by brand loyalists. This participation is increasingly crucial for small brands that do not have larger communities of customers. While generating positive information is necessary to prevent the development of Negative Cascades, firms should weigh the costs of attempting to implement this information themselves rather than using organic sources. Users of forums or communities seek them out because the information is assumed factual and helps in their decision making (Muniz and O'Guinn, 2001; Kravetz, 2007). However, users may not always know when information is intended to mislead. Brands may act as consumer benefactors by providing fact-checking and additional context that can help to incite others to respond (Scholz and Smith, 2019) and encourage a brand-favorable herd inoculation. In either scenario, brands should monitor the information sharing that takes place online, determine the accuracy of the statements, and then work with other consumers to generate a counter-response inoculation. The bystander effect may also be broken by specifically calling on an individual or group to respond (Cialdini, 1984). This feat can potentially be accomplished by

specifically calling on brand community members to assist in organizing a counter-response and prevent a herd inoculation from expanding. These implications will increase the likelihood that messages are more effectively shared and that the bystander effect is overcome.

Theoretically, these findings have implications for inoculation literature. Inoculation theory discussed how individuals could become immune to counter information (McGuire, 1968). According to Inoculation theory, individuals can build up a defense to counter-information and reinforce their views by developing a preemptive response (Banas and Rains, 2010). However, the theory does not explicitly discuss the presence of a dual-process mechanism at play and does not address the means by which messages are shared from the consumer's point of view. The findings here show that confirmatory information is much more likely to inoculate larger groups since difficulty in decision-making is reduced. Therefore, a consumer who has never heard of a brand, then sees negative information from multiple sources, may resort to an automatic system of decision-making, and join the herd's inoculated platform. However, information that disconfirms a message makes the decision-making process a little more resource consuming and thus reduces the likelihood to share a message or to create a message to be shared.

As discussed in Essay One, the real damage of a single malicious poster comes from the multiplier effect of group sharing. If building positive word of mouth can reinforce brand image, so too can negative information online tarnish brand image. By countering the Negative Cascade herd inoculation with positive information, bystanders exposed to the misleading post may choose not to join in the information sharing since other posters do not corroborate the information.

Limitations and Future Research

While the current research has novel findings with important implications, it is not without its limitations and opportunities for future research. To maintain consistency, the studies the same initial messages (discussing issues with a company removing reviews) and corresponding Negative/Positive Cascades. Therefore, future studies should focus on additional messages to determine the role of message content in driving message sharing. Additionally, the interactions took place on a single platform, Reddit, which allows greater levels of anonymity than Facebook or other communities. Future studies could review the results in communities that are more established and those that are viewed as more open/closed to new members. Study 3 focuses on a broad interpretation of the bystander effect for parsimony. Future studies should further develop the types of bystanders and the underlying influences of group cohesion, brand relationship, expertise, and perceived ambiguity in the expected response.

The treatments used the same brand – a fictional one that was created by the author – to prevent bias from affecting the results of the studies. Future research could focus on established brands in determining the strength of the bystander effect and the strength of availability cascades necessary to influence participation and information sharing. A cascade of negative information could be less daunting for established brands and may actually create its own version of the bystander effect. Future research could examine overwhelmingly positive information and the necessary effort of participants to sway opinions. The studies also only looked at Authentic/Negative Tone and Negative/No/Positive Cascades in enhancing believability, justification, and message sharing. Although other variables were tested, future studies could examine message content, brand familiarity, loyalty, and credibility of both the message and poster. Future studies can implement

biometric measures such as facial recognition and eye-tracking to triangulate the negative/positive responses that emerge.

While the primary data collection for Studies 1a and 1b consisted of an online panel purchased from Qualtrics, an M-Turk sample was used in the pretest of the survey items in the MANOVA and main data collection in Study 3. Since the study focused on whether or not a shopping situation took place online or offline, the sampling method was deemed appropriate as a wide range of consumers participate in both shopping situations. Also, the author set parameters that would only include U.S. respondents, and only those with a 95% work acceptance rate were considered for participation. While care was taken to provide a scenario appropriate for these samples, and the methods employed best practices for M-Turk (Hulland and Miller, 2018; Hulland, Baumgartner, and Smith, 2018; Sharpe Wessling, Huber, and Netzer, 2017), future research might replicate these findings using a different sampling approach

Conclusion

The purpose of this research is to start a larger narrative in understanding misleading online information sharing and the impact of herd inoculation in online communities. The “authentic” motives of individuals, albeit sometimes deceitful, can heavily sway online discourse. When these individuals actively mask intentions, confirmatory information greatly increases whether or not others continue the information sharing and strengthens the group-level inoculation. The automatic components of decision-making can be easily swayed by heuristics related to information availability. This shortcut can be misleading, and activating more cognitive resources can be difficult for not only consumers to implement but also for brands to motivate. When message-sharing is better coordinated by the brand, positive information-sharing can be emboldened, and inoculation toward negative information can be implemented.

CHAPTER IV

ESSAY THREE

Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information

Abstract

Consumers who exhibit vigilantism, those who believe that their view is right and should be shared with others, are likely to use online platforms for personal and social gain. In isolation, a misinformed message may have little effect or malicious intent. Shared beliefs that form through repeated exposure to the misinformation, or a Negative Cascade, can be devastating for brands when shared messages inoculate (immunize) other consumers against a corrective brand response. Previous research has examined the effects of so-called firestorms of information that engage consumers in large numbers. This work introduces fauxstorms as a surge of misinformation that quickly builds upon itself and gains credibility through online sharing. Brands that initiate a quarantine, by effectively isolating the vigilante from other consumers, can reduce a vigilante's ability to justify their actions to themselves and others. In doing so, a brand can build a pseudo-ally – a party that reluctantly joins ranks with the brand – to halt and prevent future misinformation creation. Study 1 finds that overt vigilante responses may halt the future sharing of misleading information. Study 2 finds that brand-enacted quarantines using Clout-based messages may act to prevent the autoinoculation of the vigilante and, thus, future misinformation sharing. By exploring these conditions, marketers can be informed about strategies to reduce autoinoculation by the individual as well as the consumer herd inoculation that spreads to others.

“If your order slips through, then lucky you, but if it's canceled [sic] and other people got in, don't grieve the poor store that just lost hundreds or thousands of dollars due to a price mistake. You're just wasting your time going down the route of trying to force them to honor a price mistake.”³

Introduction

Online communities create an anonymous habitat for those who wish to sway others, either for some unearned benefits or just for the fun of causing chaos (Binns, 2012; Suler, 2004). Brands and consumers are becoming increasingly concerned with fake or *faux* information such as the presence of post-truth politics (Roberts, 2016), truthiness, sharing information that simply *sounds* accurate, (Meddaugh, 2010; Munger, 2008), and fake information-sharing intended to cause harm (Newman, Garry, Bernstein, Kantner, and Lindsay, 2012). Although misinformation-sharing online can have negative consequences for brands, not all consumers have solely malicious intent.

Consumers may also use these online platforms to share views and organize other consumers for a personal or shared purpose (Muñiz and Schau, 2007; Saucier, Webster, Hoffman, and Strain, 2014). Vigilantism, the belief an individual has that their own views are superior and should, therefore, be spread to others, is quite common in online information-sharing (Saucier and Webster, 2010). However, online vigilantes who share misinformation in pursuit of a goal or agenda may act to justify their actions to others in their aim to gain

³ An online consumer on Reddit who describes a pricing mistake that was exploited and the attempts by the brand to address the associated responses of those who did not receive the pricing.

participation from a broader community (Schau, Muñiz, and Arnould, 2009) for a shared goal (Saucier, Webster, Hoffman, and Strain, 2014).

Although overt trolling behaviors, destructive or disruptive actions in online settings, can have devastating consequences for brands (Buckels et al. 2014), so too can vigilantes cause harm when behaving recklessly. For instance, some vigilantes may act maliciously for the sole purpose of harming another party, while others may view themselves as an advocate for other consumers. In a recent case, agrochemical giant Monsanto was attacked in 2018 by online trolls attempting to influence the local production of agricultural products (Cremer, 2018). Additionally, in 2016, a manufacturer's pricing mistake that automatically adjusted online retailers' displayed pricing caused a swarm of negative posts on Reddit directed at the unsuspecting retailers, even though many of the initial posters recognized that the retailers were not at fault⁴. These incidents show that companies of all sizes can be on the receiving end of this damaging behavior when vigilantes aim to sway the views of others.

Online vigilantes aim to take matters into their own hands and determine what is right and just (Kravetz, 2007; Muñiz and Schau, 2007) and, in doing so, aim to enhance, or promote their views to other consumers (Saucier and Webster, 2010). Online vigilantes may coordinate negative reviews to "bomb" average review scores on Yelp, TripAdvisor, Amazon, and IMDB to cause intentional harm to brands (Gonimah, 2019). Consumer-inspired negative information is troubling for over 75% of business owners who state that online reviews are essential, and one-sixth of those owners said bad reviews could potentially ruin their business (Pickard-Whitehead,

⁴ Essay One found that the Reddit forum was used to find and exploit pricing mistakes and engage in predatory shopping

2017). Brands, then, must quickly mitigate the effects of misinformation on online platforms when content is readily accessible and easily shared.

Although online community moderators can prevent some overt misconduct in forums such as Reddit, brands have little control over the creation and dissemination of these discussions (Bergstrom, 2011). Moreover, a vigilante may create a negative message that flies under a brand's online radar. The vigilantes must rationalize their intentions and determine how others will view those actions. In isolation, these individuals can build up justification for their misbehavior by attempting to defend the actions to themselves and others (Harris and Daunt, 2010). For example, a vigilante that aims to influence others (Saucier and Webster, 2010) might know that receivers viewing the message may respond unfavorably. Therefore, a vigilante may first attempt to think of preemptive retorts to persuade themselves to post negative reviews, scrutinize brands, or reinforce their viewpoint. Inoculation theory provides the theoretical lens to explain this phenomenon.

Inoculation theory suggests that individuals, when feeling that their view is correct, can become immune to counter-persuasion by building up defenses to counterviews (Pfau, 2007). Posters with vigilantism proclivities, along with such cognitive justification, begin to autoinoculate (self-inoculate) when they begin to reinforce their viewpoints to hinder a brand's response. Moreover, this process may require more mental effort to reinforce a view beyond initial judgments to more in-depth self-rationalization (Kahneman, 2011). Vigilantes not only believe in their view but also feel they are fully justified to share it with others (Saucier and Webster, 2010).

Specifically, a vigilante poster can realize a belief (“I really want the pricing, and I deserve it”) and then justify the actions (“the companies are liars, cheats, and they deserve to pay”) to create the counterargument through misleading information. Additionally, a vigilante may also feel that a company is large enough that they will not even notice any adverse effect. When vigilantes are subsequently exposed to messages that challenge their attitudes or beliefs, they are more likely to disregard them (McGuire, 1968; Compton and Pfau, 2009). Dual-process models of decision-making suggest that individuals use both automatic and analytic components when weighing choices (Kahneman, 2011). Dual-process models describe how individuals make decisions and have been applied to explain cultural norm development (Lizardo et al., 2016), trust formation (Murray et al., 2011), ethical decision-making (Haidt, 2001), and predisposition to others (Dovidio, Kawakami, and Johnson, 1997). In making decisions to share misleading information, a vigilante may be driven by both impulse and the cognitive justification for their viewpoint (Kahneman and Frederick, 2002). In isolation, these messages may have little effect, but can be damaging to brands when they begin to spread to larger consumer groups.

Consumer vigilantes, who aim to educate or influence other consumers (Larson and Denton, 2014), may determine that some degree of misbehavior is necessary for the good of their cause (Daunt and Harris, 2011; Kucuk, 2010) and the initiation of negative word-of-mouth (WOM) is warranted (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). Therefore, a vigilante may serve as the match that sparks more substantial negative information sharing online.

Firestorms, a sudden surge of negative word-of-mouth, can plague a brand after a failure (Herhausen et al. 2019), and *fauxstorms* of damaging misinformation may have similar effects

when initiated from a malicious or even misinformed vigilante – *by aiming to sway others to participate and inspiring continued group-level misinformation-sharing*. When a vigilante creates a message and shares it with a community, a critical mass can be developed through an availability cascade – *a collective belief that gains plausibility as it is repeatedly shared throughout online communication* (Kane and Webster, 2012). When a Negative Cascade forms, the initiating vigilante may become less necessary for the dissemination of a message since the cascade provides the mechanism for the misinformation to spread through herd inoculation, or consumer group immunity to a brand’s response. As a result, the inoculation from the vigilante begins to spread organically to others.

Prior literature has explored the types of responses that brands can implement to handle misinformation sharing by censoring a poster (DeKay, 2012), identifying the participant (Edstrom, 2016), attacking the information head-on (Dootson, Lings, Beatson, and Johnston, 2017; Scholz and Smith, 2019), encouraging regulations (Golf-Papez and Veer, 2017), contacting the original poster (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019), and creating an ally out of the original poster through collaboration (Wolfgang, 2018). Brands, then, have a plethora of approaches for responding to misinformation sharing, but each of these methods comes with a variety of challenges.

Since inoculation against corrective brand information requires a delay, or incubation period for persuasion to take effect (McGuire, 1964), excessive self-persuasion can backfire (Banas and Rains, 2010). For example, if a consumer is exposed to a message that is too obvious in its intention to persuade, then the inoculation process is less likely to take effect (Banas and Rains, 2010; Eisend, 2006) since an ulterior motive may be anticipated. Therefore, vigilantes that

aim to persuade themselves of a viewpoint must allow time for the position to be gradually reinforced.

Brands must also find ways to preempt potentially misleading information and prevent both full autoinoculation and herd inoculation from developing. This crucial incubation window can provide an outlet for brands to prevent the autoinoculation from fully developing. The concept of quarantine – *a state of temporary isolation* – can provide a mechanism to slow down the inoculation process. For example, a vigilante that has not developed a justification for their beliefs may not feel comfortable in sharing messages with others. This incubation window, required for effective inoculation, allows justification to form and then bolster defenses against a brand or consumer counter-response. By diminishing the autoinoculation in the incubation, or delay stage of inoculation (McGuire, 1964; Banas and Rains, 2010), a brand may reduce the vigilante’s ability to defend their actions and may thus decrease the harm caused by misinformation-sharing.

The purpose of this research is to understand how misleading information can be eliminated by first determining the role of the vigilante in halting the growth of a Negative Cascade – *a shared belief created from repeated exposure*. These individuals may feel fully justified in sharing misinformation but may also be swayed when proper communication from the brand is provided. Moreover, vigilantes may realize a mistake and quickly act to contradict their initial statement and advocate for both a brand and other consumers. By countering their own statement, vigilantes may act to halt the future spreading of herd inoculation by breaking the effect of the Negative Cascade. The following studies explore whether a vigilante can act to offset herd inoculation by overtly responding to others to gain attention. However, persuading vigilantes to contradict themselves or admit fault may require particular tact from a brand

(Baron, 1998). If a vigilante may work to halt a cascade once it has started, what can a brand do to encourage this behavior?

By determining the possibility of quarantining, or preemptively communicating with the vigilante poster of a misleading message, brands can deter the effects of autoinoculation. This quarantine is accomplished by addressing the vigilantism propensities of the individual as well as the required analytical resources of justification. In doing so, brands can break through online noise (Phan and Godes, 2018), seek out those aiming to “bomb” review scores (Gonimah, 2019), and reduce the presence of misinforming negative reviews (Pickard-Whitehead, 2017).

Two studies uncover the ability of brands to deter and eliminate misleading information before a cascade of information can take effect and spread to other posters. The implications of these two studies provide crucial insights to practitioners by providing an outlet for treatment directly at the source. By making an unlikely ally (Wolfgang, 2018), brands may employ additional resources to reinforcing positive affect and productive word-of-mouth (Muñiz and Schau, 2007). Additionally, the studies extend the boundaries of inoculation theory by addressing how the inoculation process can be reversed and the role of a quarantine in preventing the development of autoinoculation.

Background

Consumer Advocacy

An ever-growing body of literature exists that explains the role of consumers in educating, persuading and informing others (c.f. Muniz and O’guinn, 2001; Schau, Muñiz, and Arnould, 2009; Algesheimer, Dholakia, and Herrmann, 2005). When consumers advocate for a brand, they act to activate and embolden others (Wallace, Buil, and Chernatony, 2014). These

empowered consumer advocates may serve as a consumer and brand vigilante who acts to mitigate the harm of negative word-of-mouth (Muñiz and Schau, 2007).

Consumer and brand advocacy groups aim to protect consumers from misleading tactics of marketing campaigns (Holt, 2002). Campaigns that raise issues may be discussed in consumption communities where consumers compare alternatives (Stokburger-Sauer and Wiertz, 2015) and result in future brand avoidance when comparisons indicate community needs are not being met (Jayasimha, Chaudhary, and Chauhan, 2017). This advocacy for other consumers, and oneself, can also influence *negative* behaviors and information directed at brands. As such, consumer advocates can produce serious challenges for a brand when they defect from the traditional buyer/seller paradigm set by the marketplace. Just as a criminal vigilante might, a consumer vigilante may decide to take matters into their own hands.

Consumer Vigilantism

When consumers have internet-enabled access to others, some individuals may exhibit vigilante characteristics through conveying consumer advocacy motivations and attempt to protect others from companies that are deemed unethical or immoral from the poster's perspective (Kravetz, 2007; Coles and West, 2016). For instance, a consumer may actively advocate for other consumers by monitoring price fairness and positive word-of-mouth by lashing out against a brand they deem to be unfair (Wakefield and Inman, 1993; Larson and Denton, 2014). These consumers may also begin to monitor a brand's marketing communications diligently to ensure that accurate or fair information is shared by those organizations (Hsiao, Shen, and Chao, 2015).

In today's age of social media, vigilantes can leverage more substantial scale effects through social media use. For example, a small computer company, rather than Nissan Motors,

owns the domain, “Nissan.com.” The organization spotlights ongoing legal battles with the automaker and has developed a forum for consumers to express contempt for Nissan motors (Nissan.com). Moreover, the participants feel that these discussions against their “nemesis” brand are a justified and noble endeavor (Hollenbeck and Zinkhan, 2006; Kucuk, 2010). This consumer vigilantism can become even more problematic as websites, and social media discussions spread to other online platforms.

Other vigilantes can actively engage in online behavior to exploit price differences, share negative reviews, and organize group responses to confront a brand (Denegri-Knott, 2006). When vigilantes deviate too far from the norms of a productive discourse (i.e., tracking and comparing brand responses, coordinating negative reviews, and sharing unverified information), significant issues may arise. These deviant behaviors must be addressed to ensure that proper relationships are maintained (Daunt and Harris, 2011).

Consumer Deviance

While some level of consumer vigilantism inspires unsponsored communication, consumers can engage in practices that may actively harm an organization. For instance, the concept of consumer deviance includes both fraudulent and negligent components (Moschis and Cox, 1989). Fraudulent deviance typically describes consumer behaviors that are explicitly illegal and cause harm to the firm, including shoplifting, physical damage, and illegal practices (Dootson, Lings, Beatson, and Johnston, 2017). Negligent acts such as excessive returns, complaints, and harmful information-sharing can negatively influence overall profitability just as fraudulent acts can do (Fullerton and Punj, 1997; Daunt and Harris, 2011).

Whereas overtly fraudulent behavior may be more visible, negligent practices are more challenging to monitor. Vigilantes can also engage in negligent deviance, harmful yet legal

behavior, which extends beyond in-store actions to online discussions and information sharing (Binns, 2012; Coles and West, 2016). Anti-brand websites (c.f. Kucuk, 2008; Kucuk, 2010), Reddit forums ⁵, and anti-brand communities (Hollenbeck and Zinkhan, 2006) that promote anti-consumerism ideals (Krishnamurthy and Kucuk, 2009) may provide numerous outlets for a vigilante to share content with others. Many consumers may accept negative stories or personal views shared by others as fact since the receivers tend to piece together a narrative based on stories that are told by the participants (Render, 2018; Pickard-Whitehead, 2017; Kahneman, 2011). Consumers may act vigilantly by questioning company responses, examining information communicated, and comparing prior experiences from similar situations (Larson and Denton, 2014). By communicating effectively with others, vigilantes may find readily available confirmation for their views and reinforcement for their preferences (Kahneman, 2011).

Inoculation Theory

Inoculation theory suggests that individuals can develop an “immunity” from counter-information by having their existing views reinforced (McGuire, 1968). Once a vigilante builds up their own defenses against external scrutiny and reinforces their views before engaging in a deviant or aggressive behavior directed toward a brand, they can then justify, or defend, these actions to others (Inman and Zeelenberg, 2002; Daunt and Harris, 2010). Therefore, a vigilante poster may feel that an action should be taken and may then weigh the pros and cons of said action. These vigilantes, although in a more anonymous environment, still must determine if an action may draw external scrutiny (Suler, 2004; Pfau, 2009).

⁵Essay One found that Reddit forums are used to spread information to others and can be used to coordinate anti-brand responses

If a vigilante cannot rationalize sharing the misleading information or influence others without being found out, then they will not be able to reinforce their view or, thus, autoinoculate themselves. Therefore, the process will then be unlikely to affect the community as a whole through herd inoculation since the shared immunity does not have time to develop. When the vigilante develops both automatic and analytic justification, they can be more confident in sharing their message (Kahneman and Frederick, 2002).

Herd Inoculation

A vigilante's influence and persuasion in a group discussion may serve to reinforce preexisting views and ideals held by others in reference to a brand (Kravetz, 2007; Wallace, Buil, and Chernatony, 2014). Herd inoculation is defined as *the group level inoculation or immunity that forms as larger numbers of consumers inoculate against a brand's response* (Banas and Rains, 2010; Lee and Male, 2011). Therefore, the group-level belief and support that diminishes external information effectiveness can begin to take effect as other consumers join the discussion in support of a vigilante's message (Banas and Rains, 2010). In doing so, the group can maintain their similar views and spread those views to others much in the same way that herd inoculation prevents a biological threat from entering a large population (Lee and Male, 2011).

For consumer herd inoculation to occur, a large number of consumers must share the misinformation initiated by the vigilante. In doing so, a shared sentiment becomes the prevailing norm. The growing availability of the message can reinforce credibility and enhance the inoculation process in others. Prior studies suggest that both message belief and justification in future sharing is required to expand the consumer herd inoculation⁶.

⁶ The findings of Essay Two suggest that dual-process inoculation require both automatic (initial belief) and analytical (justification) components

The more a vigilante's message is shared, the more that belief can take root in the minds of others. Herd inoculation suggests that a large group of individuals can become codependent in preventing external views from breaking through and influencing the views of the group at large (Banas and Rains, 2010; Alfano, 2011). When consumer herd inoculation begins to build, dual-process models of decision-making suggest that individuals receiving the message will resort to both automatic and analytic components when weighing choices (Kahneman, 2011; Kahneman and Frederick, 2002). For instance, consumers may resort to truth default when exposed to a message that seems plausible, or posters convey guidance on recommended actions (Levine, 2016; Phan and Godes, 2018). Next, consumers then must determine whether the message can be logically supported (Kahneman and Frederick, 2002) before finally deciding to share the message or participate in the discussion (Grau and Folse, 2007)

Repeated exposure enhances the herd inoculation as other recently inoculated posters begin to share similar information or re-share the original message. In the online setting, an echo chamber effect may reinforce the views of a larger group and prevent an external brand response from permeating the group discussion (Hewett et al. 2016). A Negative Cascade provides the mechanism for herd inoculation to both strengthen and spread to others (Kane and Webster, 2012; Gaziel Yablowitz and Raban, 2016; Pfau, 2009).

Availability Cascades

Collective beliefs gain plausibility as they are shared throughout online communication (Kane and Webster, 2012) because individuals utilize the availability heuristic (i.e., a mental shortcut that relies on readily accessible or easily retrievable information) in decision-making (Kuran and Sunstein, 1998; Kahneman, 2011). Such availability cascades have been shown to influence views on health issues (Barr, 2013), reinforce extreme political viewpoints (Sun,

2012), inspire corporate regulation (Kuran and Sunstein, 1998), and reinforce biases (Alfano, 2011).

Availability cascades also enhance information credibility and plausibility (Gaziel Yablowitz and Raban, 2016). As growing numbers of consumers bolster information-sharing, the original message can be heavily reinforced as fewer and fewer counterpoints can be seen. When *faux* messages are reinforced and become more plausible, a Negative Cascade can form that reinforces the misinformation and enhances herd inoculation through repeated sharing (Kuran and Sunstein, 1998). As the misinformation spreads, more consumers exposed may then share the messages with others making it appear more credible. Thus, a prevalent Negative Cascade of misinformation may then drive future sharing (Kane and Webster, 2012) and reinforce a view (Alfano, 2011), causing a fauxstorm of misinformation that works to spread to broader groups of consumers - *a sudden discharge of misinformation that gains credibility as a Negative Cascade is created and spreads the message to larger consumer groups.*

This is especially true in communities and forums, such as Reddit, that have voting features that allow members to self-select information importance through up-votes and down-votes. When the dynamics of a group begin to lean in a particular direction, the topics, tone, and sentiments can be monopolized by homogenous views (Kane and Webster, 2012; Kuran and Sunstein, 1998). A Negative Cascade drives misinformation spreading by providing available information to influence both belief and justification in sharing the message with others (Kahneman and Frederick, 2002).

For the consumer herd inoculation process to continue, the vigilante generates a judgment in others that the message itself is believable (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000). Next, a consumer exposed to the vigilante's message must also feel that sharing the message is justifiable or defensible to others (Inman and Zeelenberg, 2002). Although a consumer may view a statement as believable, this information may not rise to the necessary threshold for future sharing. Essentially, when both automatic (initial belief) and analytic response (justification) are consistent, the consumer herd is more likely to share a message with others and strengthen the herd inoculation (Banas and Rains, 2010; Kahneman, 2011).

Hypothesis Development

Believability of the Message

For an individual to be inoculated, they must first be exposed to some type of preliminary message about a brand (Banas and Rains, 2010). Initial belief in the message sets the stage for rumination and then acceptance. An initial narrative can provide a subtle nudge by specifying an event, conveying a concern, sharing a negative view, or conveying benign intent. The believability of the message, an assessment that facts are accurate and acceptable, improves product assessments (Gürhan-Canli and Maheswaran, 2000), consumer organization (Sen, Sankar, Zeynep Gurhan-Canli, and Morwitz, 2001), and increases the perceived honesty of a source (Moore, Mowen, and Reardon 1994).

The emotional tone of a message has been shown to influence the belief in negative WOM when the overall group tone is negative (Herhausen et al. 2019). However, a negative tone in itself will not necessarily enhance the believability of content since messages lacking contextualization provide less articulate narratives (Kahneman, 2011) and detract from perceived

truthfulness (Pennebaker, 2011). Messages that convey social, personal, and ethical experiences further enhance the believability of a message (Schwarz, 2004).

By communicating information effectively, the vigilante can capitalize on truth default since people tend to assume a person is telling the truth until proven otherwise (Levine, 2014). Accordingly, when the negative tone of a message is matched with a description of the event providing context, it can drive a greater belief in the source and the message itself. For example, a vigilante that seems to have prominent and malicious motives may not have the same impact in convincing others as someone posting misinformation in a more descriptive manner, since narratives influence immersion in the content (Kahneman, 2011).

Justification in Sharing the Message

When an individual justifies their feelings or motivations, they feel more comfortable with pursuing an action. Justification in information-sharing suggests that the message that is received is acceptable and defensible (Inman and Zeelenberg, 2002). A consumer who can justify their decisions also reduces internal conflict and reinforces negative viewpoints.

Justifying a message is a fundamental component of inoculation in that the receiver must assess the information and determine its legitimacy. However, justification requires cognitive resources (Kim, Kim, and Park, 2012), and implementing these resources can be difficult in the presence of new information (Park and Cho, 2012; Kline and Wagner, 1994). When information is readily available as the basis for an opinion, this justification can be reinforced by more impulse-based decision-making (Spears, 2006). When the message received is deemed acceptable, the vigilante's message may then be shared with other parties. When the initial "gut" reaction is to believe a statement, the individual may rely on heuristics, such as the content (Meyers-Levy and Maheswaran, 2004), tone (Werle and Cuny, 2012), or quantity of messages

(Dawson and Brashers, 1996). Additionally, when a consumer deems that the actions of another are justifiable, those actions are more likely to be endorsed (Malaviya and Sivakumar, 2002).

Believability and Justification

Although the believability of a message is necessary to drive the sharing of content, it may not be sufficient. Consumers can quickly determine whether a message is worth sharing based on their personal interest in the topic (Chalkiti and Sigala, 2008), the level of negativity conveyed (Chen and Lurie, 2013), or frequency of the topic in discussions (Kahneman, 2011). Consumers may also be more likely to share a message that confirms a particular belief or preference (Lallement, Dejean, Euzéby, and Martinez, 2019).

When a consumer accepts the content as more accurate, they may be inclined to exert effort in supporting the message by providing additional supportive actions (Ball, Coelho, and Machás, 2004). When the consumer can justify sharing a message, the inoculation initiated by the vigilante begins to take hold (Meyers-Levy and Maheswaran, 2004; Kahneman, 2011; McGuire, 1964). At this point, the inoculation can disseminate throughout the consumer herd through a Negative Cascade, passed on by newly inoculated posters (Kane and Webster, 2012; Lee and Male, 2011).

A vigilante poster can provide a starting point for the decision-making of others by tapping into the more automatic components of decision-making (Kahneman, 2011). Automatic responses to an available message can make decision-making much easier with less cognitive strain when prior views or beliefs are confirmed (Evans and Stanovich, 2013). Messages that are deemed more believable appear to provide more legitimacy and trustworthiness to the content (Munnukka, Uusitalo, and Toivonen, 2016). For instance, prior studies found that an overarching negative community tone suggests the presence of shared beliefs (Herhausen et al. 2019;

(Dechêne, Stahl, Hansen, Wänke, 2010), and posters that convey a negative emotional tone while demonstrating authenticity are more believable than those sharing a negative message alone ⁷. A message that seems legitimate may be enough to generate initial belief by the consumer (Evans and Stanovich, 2013). Thus, consumers exposed to a misleading message from a vigilante can drive herd inoculation by reinforcing the views of others (Johnson, Badger, Waltermire, Snyder, and Skorupka, 2016) and can enhance the likelihood that a view is adopted (Zhu et al., 2012).

Since herd inoculation requires both automatic and analytic components, it is expected that those who believe the vigilante's message will be more likely to feel justified in sharing the message. Therefore, the likelihood to spread the vigilante's message is expected to depend on the justification of information sharing. It is also predicted that justification will mediate the relationship between the believability of the message and the likelihood of spreading the message. Therefore:

Hypothesis 1: The positive relationship between the believability of the message and the likelihood to spread the message will be mediated by the level of justification in sharing the message.

Halting Herd Inoculation

To change consumer perspectives or existing beliefs, ample evidence is required (Kahneman, 2011). In the presence of large amounts of contrary information, beliefs become increasingly difficult to change (Banas and Rains, 2010; Dawson and Brashers, 1996). Once a

⁷ Essay Two found that authentic messages conveying negative tone are more believable than messages conveying only negative tone

Negative Cascade creates a reinforcing but unobvious echo of a view through frequent exposure to a piece of misinformation, others may latch on to the message and build up a belief (Banas and Rains, 2010).

As the vigilante's message or view is spread by other receivers of the message, it then enhances the inoculation by creating a herd immunity to counter-persuasion from a brand. Thus, a vigilante may, as the cascade begins to grow organically (Kane and Webster, 2012; Kuran and Sunstein, 1998), lose control over future information-sharing. For example, a Negative Cascade that confirms the views of the original poster is more likely to be shared and less likely to lose strength since information availability is consistent with the truth default heuristic – *people assume others are truthful because dishonesty does not seem likely or there is a lack of proof that they are being misled* (Levine, 2014). However, the originator may determine that a message can no longer be defended and may even be incentivized to change their own view (i.e., through direct communication, compensation, or warnings). A brand may also convince the vigilante poster to direct responsibility to another or to an unforeseen circumstance (glitches, issues, or misunderstandings), if enough influence, proof, and justification are provided (c.f. Herhausen et al. 2019; Wolfgang, 2018; Scholz and Smith, 2019; Dootson, Lings, Beatson, and Johnston, 2017).

Not all vigilantes operate with mal intent; some focus on acting as a benefactor for other consumers (Larson and Denton, 2014), aim for fairness (Coles and West, 2016), or initiate price comparisons (Wakefield and Inman, 1993) while taking actions that a brand may find to be undesirable (Moschis and Cox, 1989). Thus, when handling a firestorm of legitimate consumer information, brands may contact the poster and have them remove a negative review or post (Herhausen et al., 2019). In doing so, a firestorm, over time, may dissipate. However, a

fauxstorm may require more effort from the original vigilante poster since the misinformation may be challenging to confront or correct.

If a vigilante poster determines that additional actions are needed to offset the misinformation-sharing due to new information, attention should be provided to gaining the necessary exposure to counteract information manipulation (Dawson and Brashers, 1996; McCornack, Levine, Solowczuk, Torres and Campbell, 1992). The overall importance placed on the message may act to solidify the inoculation process (Pfau, 1997). So, a strong statement from the originator may be required to halt the process. Therefore, the original poster must use logical extensions to refute their own original counterargument while tapping into the beliefs, attitudes, and preferences of the consumer herd. The original poster may slow the herd inoculation by preventing the Negative Cascade from continuing to spread.

Vigilante Responses

Healthy consumer interactions become increasingly important in maintaining information integrity online for both consumers (Larson and Denton, 2014) and brands (Kennedy, Lawton, and Plumlee, 2002). When interactions deviate from accepted norms, both brands and consumers may be required to make corrections (Holman and Lay, 2019; Kennedy et al. 2002), update information (Kalyanam, McIntyre, and Masonis, 2007; Shaw-Ching Liu, Sudharshan, and Hamer, 2000), establish symmetry in information sharing (Friedmann, 2019), and share new information when posting online (Pongsakornrungrungsilp and Schroeder, 2011).

Vigilant consumers have a plethora of opportunities to update and share new information with others as it becomes available. However, they must also weigh the choice of response with the available platforms. For instance, a vigilante may aim to correct a previous statement, if proven wrong, by simply posting a message to community members. In doing so, the vigilante

may feel a discreet response – *an unobtrusive response within the community to provide new information or corrections*– will suffice (c.f. Muniz and O’Guinn, 2001). However, the vigilante may also feel the need to bring additional attention to the correction and aim to provide an overt response – *a response to a community aimed at drawing explicit attention to new information or corrections* – or more prominent update to the shared misinformation (c.f. Larson and Denton, 2014).

Discreet Vigilante Responses

Individuals who complain online are often asked by the brand to remove a negative review (Herhausen et al., 2019), but the overall negative effect may still exist for some time before a firestorm subsides (Herhausen et al., 2019; Khamitov, Grégoire, and Suri, 2019). Censoring the content itself by the brand or the vigilante poster still might cause concerns from others who have viewed the original message (DeKay, 2012). Therefore, it is crucial for the vigilante that posts misleading information to not only remove but also correct the message to halt future sharing.

Just as a newspaper that issues a rejoinder may make updates within a future edition, so too can online posters effectively edit previous messages. For example, studies have recommended that companies update social media information (Beukeboom, Kerkhof, and de Vries, 2015; Ayu and Abrizah, 2011) and update information in a consistent location (Ashley and Tuten, 2015) to reach others through multiple internal exchanges (Brown, Broderick, and Lee, 2007). Additionally, a vigilante consumer may aim to communicate directly to maintain existing current relationship ties created through previous communication to a currently captive audience (Underwood, Kerlin, and Farrington-Flint, 2014; Hausman, Kabadayi, and Price, 2014).

A forum may be created solely for the benefits of its members (Muniz and O'Guinn, 2001), and consumers may prefer updates to communications be made within the same platform for further discussions (Fournier and Lee, 2009) to encourage collaboration (Schau, Muñiz, and Arnould, 2009). For example, in an online setting, Reddit allows consumer moderators to post updates within the discussion itself to continue a conversation. The up-vote and down-vote features then allow future posters to see the new information that is created and then shared by the original poster. Therefore, a vigilante and the community may choose to continue the discussion in the forum.

Consumers may also accept information from others who communicate in a consistent manner (Ludwig, Ruyter, and Friedman, 2013) and encourage internal cooperation (Schau et al., 2009) to discuss community topics (Fournier and Lee, 2009). When the vigilante decides to provide a response that challenges their original post, they may choose to post in the same forum to be consistent with community norms (Algesheimer, Dholakia, and Herrmann, 2005) and choose a more discreet response that continues the discussion rather than redirecting it elsewhere. Consumers may also feel that an internal forum rebuttal is necessary to maintain community norms (Algesheimer, Dholakia, and Herrmann, 2005). Consumers communicating effectively within a community allow updates to be productive and concise (Gallois and Giles, 2015). Consumers may also seek out and accept information from others who communicate in a similar manner (Ludwig, Ruyter, and Friedman, 2013) and may find current communication channels to be adequate for new information sharing (Daugherty and Hoffman, 2014). Therefore, a vigilante may choose to post a discreet response within the group discussion itself.

The posters receiving the discreet message may have difficulty in justifying the sharing of the vigilante's previous message when the new message challenges the original content.

Thus, the meaning of the original message and its perceived accuracy, in light of the new information, is in conflict (Dawson and Brashers, 1996; Levine, 2016). Therefore, it is expected that when the vigilante corrects their original message in the discussion, they can attenuate a persistent Negative Cascade.

Hypothesis 2: The mediated relationship between the believability of the message and the likelihood to spread the message through justification will be weakened by the presence of a discreet contradictory vigilante response.

Overt Vigilante Responses

Although a discreet message can potentially reduce the harm caused by a misleading message, consumers are bombarded with content on a daily basis, and such messages may go unnoticed (Kahneman, 2011; Kahneman and Frederick, 2002). Likewise, the vigilante may be able to significantly reduce future misinformation sharing by attempting to discredit the original statement explicitly and reduce their original refutation (Banas and Rains, 2010; Pfau, 2009). Therefore, messages that are more overt may be required to break through the noise (Phan and Godes, 2018). In essence, a vigilante engaging in negligent deviance may potentially halt future harm when a corrective message effectively reaches others (Phan and Godes, 2018).

This may require the vigilante to break forum or community norms to direct attention to a new discussion (Algesheimer et al. 2005). These individuals may purposely take a personal risk by confronting an informational mistake or stating that their original message was blatantly false (Baron, 1998; Leibel, 1991). Since vigilantes may aim to act as consumer advocates, they may seek out new information and be explicit in the sharing of new information. For example, vigilantes that take overt action would include a person posting directly about a brand (Kravetz, 2007), creating a new discussion to bring attention to unfair practices (Wakefield and Inman,

1993), or ensuring positive information is shared about a brand while contradicting a statement found to be inaccurate (Larson and Denton, 2014). Additionally, a vigilante may aim to broadcast an update to reach a large audience (Underwood, Kerlin, and Farrington-Flint, 2014) and aim to extend the social reach of the information (Hausman, Kabadayi, and Price, 2014).

The proper placement of these vigilante messages is paramount for others to take notice of (Lewis et al., 2013). Because a more tangible message is more memorable (Parrott et al., 2008), the message content (Kolyesnikova, Sullivan Dodd, and Callison, 2011), explicit acceptance of fault (Rosenbaum, Kuntze, and Wooldridge, 2011; Baron 1998), and exoneration of the accused (Carrigan and Attalla, 2001), can sway the sentiments of a larger group. Moreover, conveying additional action in the creation of a corrective message reduces the amount of effort required by the receiver in interpreting the message (Ball, Coelho, and Machás, 2004). Additionally, consumers are more likely to transmit negative information about another party and generate positive information about themselves (DeAngelis et al. 2012). Thus, posts that generate negative information may be more noticeable.

When a vigilante purposely calls attention to an informational update, they bring additional focus and thus show their effort in accommodating the information to others (Giles, Coupland, and Coupland, 1991). Additionally, a vigilante can create group cohesion when effectively advocating for other consumers through new information (Liu, Xie, and Zhang, 2019; Forman, Ghose, and Wiesenfeld, 2008). Therefore, a Negative Cascade can be attenuated in such situations. However challenging to motivate, consumers who admit fault and attempt to address others overtly and explicitly can attenuate the harmful effects of a Negative Cascade. Therefore:

Hypothesis 3: The mediated relationship between the believability of the message and the likelihood to spread the message through justification will be weakened by the presence of an overt contradictory vigilante response.

The hypothesized relationships are visualized in Figure 4.1.

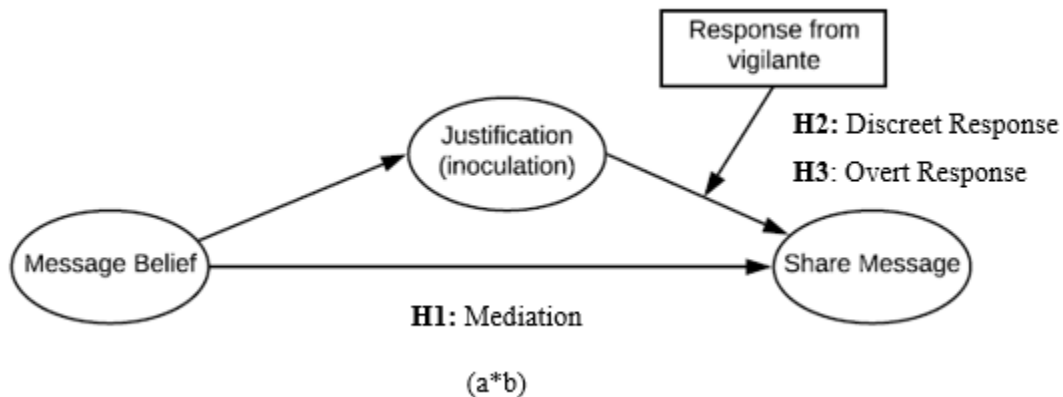


Figure 4.1 Study 1 Model

Study 1: Halting the Herd Inoculation

Method

Study 1 implements an experimental design research methodology to test the hypothesized relationships. A total of 225 participants, mean age of 39 years, and 54.7% male, were recruited through an online U.S. consumer panel purchased from M-Turk. M-Turk was deemed appropriate since the scenario focused on general knowledge of online shopping and online communication. 68.2% of the respondents indicated that they purchase either frequently or exclusively online, use social media at least daily (75.8%), and participate actively in online discussions (56.5%). (See Table 4.1).

Table 4.1 Study 1 Sample Description

Category	Total Collection (223)
Age	
18-20	2 (1%)
21-30	69 (31%)
31-40	83 (37%)
41-50	32 (14%)
51-60	15 (7%)
61 and over	22 (10%)
Gender	
Male	122 (54.7%)
Female	101 (45.3%)
Social Media Usage	
Never	4 (1%)
Monthly	14 (6%)
Weekly	36 (16%)
Daily	138 (62%)
Hourly	31 (14%)
Social Media Discussion Activity	
Not very active	14 (6%)
Not active	44 (20%)
Neither active nor inactive	39 (18%)
Active	96 (43%)
Very active	30 (13%)
Online Shopping Frequency	
Never	2 (0%)
Rarely	11 (5%)
Sometimes	58 (26%)
Often	115 (52%)
Always	37 (17%)

Participants were given a brief overview of the study and asked for their consent. Once participants agreed, they were given instructions and were shown a misinformed message created about a fictitious brand, New Latitudes. The negative message was also displayed with a corresponding Negative Cascade that showed a Reddit forum with numerous respondents criticizing the brand for deleting negative reviews. The respondents were then randomly shown

one of three conditions (no response from the vigilante, a discreet response from the vigilante within the current forum, or an overt response from the vigilante that aims to gain the forum's attention).

Each participant was told about New Latitudes and then shown the message from the vigilante and the corresponding Negative Cascade Treatment (See Appendix C). Then, they read a scenario in which the vigilante did not respond, responded discreetly within the forum and admitted fault, or responded overtly by starting a new discussion to admit fault. A manipulation check indicated that respondents saw the overt vigilante response as a more noticeable correction of the original message $t = 5.92, p < .01$ ($M_{\text{Overt}} = 5.98, M_{\text{Dis}} = 4.18$). Care was also taken to ensure that the Negative Cascade was in place by requiring respondents to respond to 1-7 scale items that reflected the negativity of forum tone. Respondents indicated that the forum was overwhelmingly negative toward New Latitudes (6.03), primarily against New Latitudes (6.01), and in agreement with the vigilante poster (5.86).

A qualitative reading check question was asked about halfway through the survey to ensure participants were paying attention; two participants who failed the reading check question were excluded from the study, and their responses were not used. The two respondents were removed using listwise deletion (Hair et al., 2010). All the thought listings were reviewed for any abnormalities, and no apparent problems were evident.

Participants were then asked to answer survey questions regarding the believability of the message ($\alpha = .91$; Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000), their justification in sharing the message ($\alpha = .94$, Inman and Zeelenberg, 2002), and their likelihood to spread the message ($\alpha = .95$, Lee and Ma, 2012). All constructs, items, and reliabilities are available in Table 4.2.

Table 4.2 CFA Results, t-values, and Cronbach's α

Scales	α	Standardized Factor Loadings	t-values
Believability- (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000)	0.91		
Highly believable		0.72	13.62
Absolutely true		0.81	17.04
Totally acceptable		0.74	14.33
Very credible		0.93	*
Completely trustworthy		0.87	19.82
The justifiability of the decision- (adapted from Inman and Zeelenberg, 2002)	0.94		
Strongly Justifiable		0.85	17.94
Very Logical		0.93	22.19
Acceptable		0.94	22.43
Reasonable		0.89	*
Likelihood to Spread the Message (Lee and Ma, 2012)	0.95		
I would share the message online		0.94	38.04
I would share the message when discussing the company		0.90	23.40
I would spread the information to others		0.91	26.85

Note: $\chi^2 = 139.06$, $df = 51$, $p < 0.01$, $\chi^2/df = 2.72$, incremental fit index (IFI) = 0.97, Tucker-Lewis Index (TLI) = 0.96, comparative fit index (CFI) = 0.97, root mean square error of approximation (RMSEA) = 0.08. All factor loadings have a p-value < 0.01 * denotes a constrained relationship to 1.00 for identification

A confirmatory factor analysis was performed using AMOS 26 to assess the unidimensionality, convergent validity, and discriminant validity of the latent constructs. The results of the analysis indicated an acceptable fit (Marsh, Hau, and Wen, 2004) of the model to the data ($\chi^2 = 139.06$, $df = 51$, CFI = .97, TLI = .96 IFI = .97, RMSEA = .08). Table 4.2 shows a complete list of results from the CFA, along with Cronbach's α for each construct.

Discriminant validity was assessed among the constructs using Fornell and Larcker's (1981) criterion. All of the constructs achieved discriminant validity as the average variance extracted for each construct was higher than the squared correlations between any pairs of constructs (see Table 4.3 for the shared variance between constructs and AVEs). The potential for common method variance was addressed by introducing a latent common method factor

(Podsakoff, MacKenzie, Lee, and Podsakoff, 2003) to ensure that the measures are adequately represented. The results indicated that the presence of the latent common method factor did not significantly change the model fit ($\Delta\chi^2/df1 < 1$). Therefore, common method bias does not appear to be a significant concern.

Table 4.3 Means, Standard Deviations, AVEs, and Shared Variance

	Mean	SD	1	2	3
Believability	5.06	1.23	0.67		
Justifiability of the Decision	5.25	1.47	0.57	0.81	
Likelihood to Spread the Message	4.24	1.81	0.18	0.16	0.86

Note: Average Variance Extracted is listed in the diagonals of the table

Results

Mediation

The bootstrap method with 5,000 samples and a 95% confidence interval was used to calculate the mediating relationships (Hayes, 2018). The results of the mediation bootstrap analysis conducted using PROCESS Model 4 are found in Table 4.4.

Table 4.4 Hypothesis Results

	Hypothesized Relationship	Unstandardized Estimates	t-Values	Indirect Effect	CI: Low	CI: High	Hypothesis Supported?
Mediation	H1: Believable=>Justify=>Share Message			0.24	0.01	0.46	Yes
Index of Moderated Mediation	H2: Believable=> Justify*Discreet=> Share Message	-0.32	-1.82	-0.29	-0.66	0.05	No
Index of Moderated Mediation	H3: Believable =>Justify*Overt=> Spread Message	-0.56	-2.99*	-0.51	-0.94	-0.15	Yes

Believable => Justifiable=> Spreading Message

The a-path (0.89, $t = 17.02$, $p < 0.01$) from the message type to believability to justification and b-path (0.54, $t = 3.47$, $p < 0.01$) from justification to the likelihood to spread the message are significant. The c-path is significant (0.37, $t = 2.82$, $p < 0.01$). Additionally, the indirect effect (0.24, CI Low = 0.01, CI High = 0.46) is also significant. Therefore, Hypothesis 1 is supported.

Moderated Mediation

Believable =>Justification*Vigilante Response =>Share the Message

The 5,000 bootstraps with a 95% confidence interval used in the mediation analysis were also employed when calculating the moderated mediation relationships using PROCESS Model 14 (Hayes, 2018; Zhao et al., 2010). The relationship between believability to the likelihood of spreading the message through justification was reviewed once again by introducing a response from the vigilante (0/Discreet) (/Overt). The interaction effect for the justification in sharing the message*Discreet vigilante was non-significant (-0.32 , $t = -1.82$, $p = .07$). Thus, Hypothesis 2 is not supported. Although community norms might dictate that messages be shared within the common community thread (Algesheimer, Dholakia, and Herrmann, 2005), the discreet message is not powerful enough to offset a Negative Cascade that has been created.

The interaction effect for the justification in sharing the message*Overt vigilante response, however, was significant (-0.56 , $t = -2.99$, $p < 0.01$). Additionally, the index of moderated mediation (-0.51 , CI Lower = -0.89 , CI Upper = -0.15) indicates that the presence of an overt message from the vigilante does moderate the indirect effect on the likelihood of spreading the message. The spotlight analysis provided Johnson-Neyman points (Spiller et al., 2013; Johnson and Neyman, 1936). Johnson-Neyman points were found when justification is less than 1.43 ($p < .05$) and greater than 4.98 ($p < .05$). This finding suggests that when justification is 4.98

or greater, the interaction becomes negative. Therefore, Hypothesis 3 is supported. Consumers exposed to an overt response from the vigilante poster are significantly less likely to share the original message moving forward. Those who were not exposed to a response from the vigilante were significantly more likely to spread the message they received, $F = 4.51$, $p < .01$, ($M_{No} = 4.74$, $M_{Dis} = 4.03$, $M_{Overt} = 3.94$). This finding suggests that individuals who begin a cascade may also act as effective agents or allies to reduce the harm of future information-sharing.

A post hoc analysis was completed to also provide additional insight and context to the respondents' views of the vigilante responses. The treatments differed significantly on the perceived genuineness – *the view that the interaction seems natural* – of the vigilante, $F = 10.10$, $p < .01$, ($M_{No} = 5.47$, $M_{Dis} = 4.97$, $M_{Overt} = 4.26$). The post hoc Least Significant Difference (LSD) post hoc analysis found that the overt response made the vigilante poster seem significantly less genuine than the discreet response or no response at all. However, the original truth default effect (Levine, 2014) appears to be broken since the respondents rated the overt vigilante response as seeming more truthful, $F = 9.55$, $p < .01$, ($M_{No} = 5.00$, $M_{Dis} = 4.19$, $M_{Overt} = 5.48$).

This outcome is consistent with findings in previous studies suggesting that individuals are reluctant to admit fault blatantly (c.f. Rosenbaum, Kuntze, and Wooldridge, 2011; Baron 1998), and respondents shared this sentiment, but also appreciated when vigilantes did so. Therefore, although the overt response reduces the likelihood of sharing the original message with others, it may raise red flags for others and may even reduce the perceived credibility of the source if a brand does not reinforce the goodwill created by the vigilante.

Discussion

The results from Study 1 support H1 and H3, and the lack of support for H2 is very informative. The findings indicate that the vigilante can have a crucial impact in halting a Negative Cascade of information. When the vigilante posts a more overt response and aims to gain the attention of the consumer herd, it reduces the other posters' ability to justify sharing the vigilante's original message. Additionally, a discreet response from the vigilante to the community (i.e., replying somewhere within the forum) fails to reduce the justification of others in sharing a vigilante's original message. Consumer responses are effective in halting the cascade when the vigilante is overt in correcting a previously misleading statement.

When exposed to this overt vigilante response, those viewing the forum are less likely to share the original message with others. However, the overt message from the vigilante may seem unusual and raise additional questions from the community members. This finding suggests that individuals may also be aware of this perspective and choose not to post overt contradictions of messages since others may view them less favorably. Admitting a mistake is difficult, and the prospect of losing face can reduce the likelihood of admitting guilt (Baron, 1998; Leibel, 1991). Respondents stated that it is more likely that a vigilante poster would not post a response, especially one that directly contradicts themselves. These findings suggest that the vigilante can diminish a Negative Cascade of information and the resulting herd inoculation if enough effort is exhibited in doing so.

Although a vigilante can be useful in halting herd inoculation and misinformation spreading, they may not always be motivated to do so since these actions require admitting guilt (Baron, 1998) and may attract scrutiny from the community members (Schau, Muñiz, and

Arnould, 2009). However, brands may communicate both effectively and preemptively to successfully prevent a cascade from forming in the first place.

Since inoculation requires a delay for persuasion can take effect (McGuire, 1964) and over persuasion can backfire when defenses are hindered by suspicion (Banas and Rains, 2010), a vigilante may not be able to justify their actions immediately. The incubation period gives a brief window to confront a misinformed view (McGuire, 1964), and brands may have an opportunity to halt misinformation before it has time to spread throughout the consumer herd. The concept of quarantine is introduced as a mechanism to slow down the inoculation process by attenuating the autoinoculation in the incubation, or delay, stage (McGuire, 1964).

Study 2 explores the brand's response in communicating directly with the vigilante as a means to prevent the initial creation of a message, highlighting the challenges faced by brands in influencing vigilantes to reverse their stance. The goal of this study is to understand the response that a brand can initiate to hinder a vigilante's ability to autoinoculate and reinforce their views. In doing so, the vigilante will be less likely to share the misinformation that can cause a Negative Cascade. Quarantines – *a state of temporary isolation* – act to prevent an infectious illness from spreading by keeping those infected separated from others (Follett and Neven, 2006).

Once a quarantine window has been met, individuals may be deemed safe to interact with others once it is determined that they are not infected (Fuhrmann, 2017). Similar inferences can be made when developing a quarantine to halt the autoinoculation of a vigilante by preventing them from persuading themselves and other consumers. For example, a vigilante that has not developed a justification for their beliefs may not feel comfortable in sharing messages with others. Therefore, inoculation requires an incubation period for views to be reinforced (McGuire, 1964).

The respondents in Study 2 are exposed to a situation in which they are in a position to exploit a pricing mistake online by sharing self-created misinformation. They are shown a number of brand quarantine responses that aim to offset the autoinoculation that takes place when a vigilante can justify sharing misleading information. The effects will then be compared to determine the most effective brand quarantine message.

Misinformed Vigilantes

Consumer communities and brand communities can advocate for a brand offering to protect loved brands by actively disputing negative information and sharing positive impressions linked to the brand (Kravetz, 2007; Wallace, Buil, and Chernatony, 2014). Additionally, vigilantes who advocate for others may build toward an active collaboration between a brand and diverse communities dedicated to purchasing decisions (Fournier and Lee, 2009; Kravetz, 2007). By accessing the shared knowledge of a vibrant community, vigilantes expose others to an abundance of both positive and negative information (Muñiz and Schau, 2007; Larson and Denton, 2014). However, the same proactivity in communication may be at play when misinformation begins to spread (Schivinski and Dabrowski, 2016).

Vigilantes may be misinformed and behave recklessly by sharing information that confirms their view without properly verifying credence or accuracy of a message before sharing with others (Levetin, 2016). Although not blatantly malicious, these actions can cause similar harm to those who are engaging in purposely malicious behavior. These vigilantes may act as active advocates for other consumers by diligently disputing statements or online information they feel is fraudulent (Wakefield and Inman, 1993; Larson and Denton, 2014), but they may be misguided in determining the accuracy of the information or the intentions of other posters. When these consumers pool informational resources, shared insights can act to enhance the

benefits of community members (Schau, Muñiz, and Arnould, 2009; Wirtz et al., 2013). However, vigilantes may be overloaded with conflicting information and may make bad decisions (Levetin, 2016). When a vigilante believes misinformation or determines their own misguided view to be accurate, they may then use their platform and influence to advocate for a misleading cause.

Quarantining the Vigilante

A vigilante that begins the autoinoculation process first determines that their view has merit and is reinforced (McGuire, 1968), whether it be through self-interest or through a perceived greater good for themselves or others (Saucier and Webster, 2010). Physiologically, vaccines work by introducing a weak form of a virus so that the immune system can learn to fight it off (McGuire, 1968). So, too, can vigilantes can also build up a defense against a persuasive brand response.

By identifying a potential threat early, the individual can protect a belief or attitude (Banas and Rains, 2010; Compton and Ivanov, 2012). For instance, inoculation theory suggests that an individual, or entity, can preemptively think of ways that a brand or company may respond and then build a defense against that action (Banas and Rains, 2010). These individuals begin to self-inoculate or autoinoculate when they start to justify their intent and goals, but this process takes time to take effect.

Once the vigilante poster has begun to autoinoculate by justifying their position to themselves, an incubation period (delay period) is required before autoinoculation, and the corresponding herd inoculation takes hold. If the vigilante offers too much supporting information too quickly, others may become aware that the originator of the message is attempting to persuade them overtly and begin to question the validity of all statements from the

originator (McGuire, 1964). For example, too much information makes the vigilante appear that they are trying too hard to reinforce an argument. Therefore, the vigilante may experience similar challenges when attempting to build up their own justification for their viewpoint.

As inoculation theory suggests, too much, or too little (McGuire, 1964), of a delay in persuasion can dilute the impact of the inoculation (Banas and Rains, 2010). It is during this incubation window that the brand has an opportunity to preempt the misinformation by going directly to the vigilante before the consumer herd inoculation begins. Quarantines prevent an infectious illness from spreading by preventing those exposed to illness from coming into contact with others (Follett and Neven, 2006). After a successful quarantine, individuals may be deemed safe to interact with others (Fuhrmann, 2017). This study explores the means by which a brand may act to quarantine vigilantes before they have the opportunity to mislead themselves and others.

Vigilantism Propensities

The concept of vigilantism, prompted by a need to transfer opinions to others (Saucier and Webster, 2010), can be used to describe actions that consumers take to educate or inform others (Kravetz, 2007). Those with a propensity for vigilantism can drive online communication since online platforms provide an ability to reinforce and share with others both for benevolent and malevolent intent. For instance, consumers may ensure that the information posted in online reviews is accurate (Larson and Denton, 2014) and monitor pricing information and policies (Wakefield and Inman, 1993). Potential brand damage can result when this vigilantism is prompted by a need to share opinions when a vigilante forgoes fact-checking, and posts without regard for accuracy (Saucier and Webster, 2010), in effect, misinforming others. This pursuit is marked by the view that the consumer's opinion is an accurate view, and it is their responsibility

to “inform” others at all costs (Saucier et al., 2014). In these situations, a vigilante may be driven to use the tools available online to share a misleading message with others if it helps them accomplish a goal.

Justification for a Decision

A vigilante must justify actions to themselves and the group, and this process requires that their actions are defensible to others (Inman and Zeelenberg, 2002). When an individual explains their feelings or motives, the message receiver can feel more comfortable engaging in activities that may be deemed undesirable by others. Individuals aim to justify their actions to reduce internal conflict (Kim, Kim, and Park, 2012). Cognitive justification is a fundamental component of inoculation in that an individual must assess and justify the presence of an external informational threat from the brand (Kahneman, 2011; McGuire, 1968). Thus, cognitive justification is expected to influence a vigilante’s ability to inoculate themselves. The justification of one’s actions helps a vigilante to maintain positive affect while twisting the truth (Schweitzer and Hsee, 2002), making purchases (Okada, 2005) accepting prices (Choi, Li, Rangan, Chatterjee, and Singh, 2014), self-persuading (Bernritter, van Ooijen, and Müller, 2017), and sharing online C2C information (O’Sullivan, 2015). Additionally, consumers aim to justify actions that support their own views (Malaviya and Sivakumar, 2002). Therefore, justification is essential for vigilantes when determining to share a misleading message with others and thus autoinoculating themselves. Therefore:

Hypothesis 4: A vigilante’s justification for their actions will strengthen the positive relationship between their vigilantism and the likelihood of sharing self-created misleading information.

Implementing a Quarantine

Quarantines prevent an illness from spreading by isolating an individual from others for a period of time (Follett and Neven, 2006). After a successful quarantine, individuals may be deemed safe to interact with others after a threat is eliminated (Fuhrmann, 2017). Although typically discussed in biomedical research, quarantine strategies have been implemented to prevent the spreading of malware (Moore, Shannon, Voelker, and Savage, 2003), protecting online consumer data (Zou, Gong, and Towsley, 2003) and reducing commerce (Sato, 2009). Moreover, some have suggested that quarantines can act as theoretical self-imposed isolation from other parties (Hartnett, 2010).

One brand strategy that might attenuate the adverse effects of misinformation is to create an ally out of a disruptive consumer (Wolfgang, 2018). For instance, brands have contemplated the benefits of confronting the challenges of attacking misinformation (Dootson, Lings, Beatson, and Johnston, 2017; Scholz and Smith, 2019) while also weighing the online privacy concerns of identifying a poster (Edstrong, 2016). Therefore, some brands have determined that more proactive and collaborative measures may be fruitful (Wolfgang, 2018). Since the literature suggests that contacting the original poster (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019) after a service failure complaint and creating an ally through collaboration (Wolfgang, 2018) provide crucial assistance to those battling firestorms and trolling behavior, these strategies can aid in the development of a brand quarantine. Moreover, these approaches can help a brand create a defense against negative consumer autoinoculation and the creation of a fauxstorm by quarantining the vigilante before misinformation is shared.

Since vigilantes, themselves, autoinoculate (McGuire, 1968)⁸, or self-justify actions that may cause harm (Kahneman, 2011), the process too may start in isolation away from external parties. Additionally, sharing a self-created message can be driven by the vigilante's individual automatic proclivities and the associated logic for their intentions to share a piece of information (Dovidio, Kawakami, and Johnson, 1997). However, this process might provide opportunities for brands since individuals with negative propensities still need to justify their actions. This necessity provides marketers with a quarantine window to offset a negative defense before the information is shared. The vigilante may need to allow the justification to set in and further develop (i.e., delay) (McGuire, 1964; Banas and Rains, 2010). This incubation period can provide the necessary window for a firm to intercede and offset autoinoculation. For instance, too much of a delay can dilute the impact of the inoculation for a marketing communication campaign by allowing counterarguments from a competitor, customer, or another party to gain prominence (Banas and Rains, 2010).

The brand can attempt to break the autoinoculation by slowing down the vigilante in sharing the message through an effective quarantine period. Since the vigilante may not have the same exposure to a collective system of belief (as the message has not been shared yet), they may be more susceptible to responses from a brand that offset the justification for the misleading actions. Therefore, a well-timed preemptive response from the brand can act to reduce the likelihood that a misleading message is created and shared with others (See Figure 4.2).

Hypothesis 5: The moderating effect of justification on vigilantism will be moderated by the presence of a brand response, reducing the likelihood to share self-created misleading information.

⁸ Essay One found that posters autoinoculate through a dual-process decision-making model

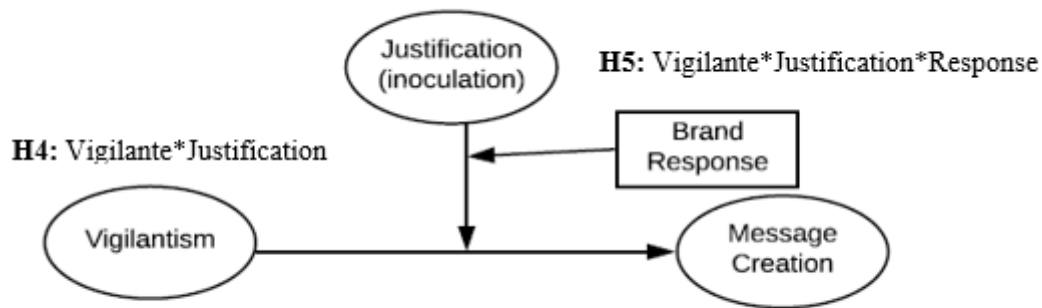


Figure 4.2 Study 2 Model

Study 2

Brands have a number of options for determining the content of messages to communicate with consumers. Therefore, additional context is needed to determine the most effective brand response to quarantine the vigilante before misinformation can spread. For this study, the brand responses are created using Linguistic Inquiry and Word Count (LIWC) to convey corresponding Analytical thinking, Authenticity, Clout, and Emotional Tone. LIWC software allows a researcher to select a specific written text type and then compare the selected text to qualitative and quantitative baselines that have been established through numerous algorithms and analyses by expert judges (Pennebaker, 2011). Additionally, LIWC’s dictionary is updated regularly and now consists of over 5000 words (Pennebaker, 2019). The software has been implemented to address marketing phenomena such as feedback loops (Hewett, Rand, Rust, and van Heerde, 2016), viral marketing (Berger and Milkman, 2012), online reviews (Ludwig, De Ruyter, Friedman, Brüggem, Wetzels, and Pfann, 2013), and negative WOM (Herhausen, Ludwig, Grewal, Wulf, and Schoegel, 2019). LIWC’s summary variables provided a framework

for the scenario and aided the researcher in developing a brand message high in each dimension. An ANOVA assessed the most effective brand responses and those that are deemed most genuine and realistic – hence believable –by participants.

The respondents were exposed to an online shopping situation in which they felt sharing misleading information might help them take advantage of a pricing mistake. The respondents were then exposed to one of the brand responses based on four LIWC treatments (Analytical, Clout, Authentic, or Negative Emotional Tone). The respondents were asked to rate the overall perceived genuineness of the response– *the view that the interaction seems natural* – with four items using 7-point semantic differential endpoints (Zhu and Meyers-Levy, 2009) and realism of the brand response consisting of one item 1-7 “very unrealistic to very realistic.” This knowledge was essential in determining which brand responses are most effective in initiating the vigilante quarantine.

LIWC Treatments

To test the brand responses, the author developed response treatments based on the summary variables from LIWC that correspond to messages that a brand might use to respond to a consumer vigilante. The four summary variables are discussed below: Analytical thinking, Clout, Authenticity, and Emotional Tone.

The Analytical thinking summary variable is a factor-analytically derived dimension based on eight cognitive-function word dimensions (c.f. Pennebaker, Chung, Frazee, Lavergne, and Beaver 2014). The analytic summary dimension addresses the level to which people use words that suggest formal, logical, and hierarchical thinking patterns (Pennebaker, 2019; Pennebaker, Chung, Frazee, Lavergne, and Beaver, 2014). In online settings, brands may aim to lead with facts to address a vigilante.

The Clout summary variable refers to the relative social status, confidence, or leadership that people display through their writing (Pennebaker, 2019). The algorithm was developed based on the results from a series of studies in which people interact with one another (Kacewicz, Pennebaker, Davis, Jeon, and Graesser, 2013). Brands may convey a certain level of importance or status based on their postings. Brands also attempt to sway others by conveying leadership, confidence, and collaboration.

Posts that appear to be Authentic can reinforce the message that is being conveyed (Pennebaker, 2019). Newman, Pennebaker, Berry, and Richards (2003) developed the algorithm for Authenticity through a series of studies in which people were convinced to be honest or deceptive. Additionally, Pennebaker (2011) contributed to the creation of the algorithm through published summaries of deception studies. When people reveal themselves authentically or honestly, they are perceived as more personal, humble, and vulnerable (Pennebaker, 2011). A brand may also speak from a place of vulnerability to reduce the vigilante's justification in their actions.

Finally, the Emotional Tone summary variables are determined for both the positive and negative emotional dimensions (Cohn, Mehl, and Pennebaker, 2004). Pennebaker (2011) states that the algorithm is built so that the higher the number (i.e., 100), the more positive the tone. Scores below (50) indicate a more negative Emotional Tone (Cohn, Mehl, and Pennebaker, 2004). This measure addresses the raw emotion that is conveyed in posts and the brand's ability to use a similar tone as those posting online.

Brand Response

A total of 300 participants that represent an average online shopper were recruited from an M-Turk panel. Four respondents failed attention checks and were excluded from the analysis.

Participants were exposed to a scenario that provided a shopping situation in which a consumer has the opportunity to exploit a pricing mistake. The scenario manipulates the resulting justification in sharing the information with consumers online by stating that it would increase the likelihood that the mistaken pricing would be honored for the reader if others were complaining. After exposure to the condition, respondents completed an online questionnaire (via Qualtrics). The sample consisted of 59.6% female respondents with a mean age of 38.61 years. 66% of the respondents indicated that they purchase either frequently or exclusively online, use social media at least daily (71%), and participate actively in online discussions (48%). (See Table 4.5).

Table 4.5 Study 2 Sample Characteristics

Category	Total Collection (296)
Age	
18-20	4 (1%)
21-30	93 (31%)
31-40	97 (33%)
41-50	47 (16%)
51-60	31 (11%)
61 and over	24 (8%)
Gender	
Male	119 (40.2%)
Female	177 (59.8%)
Social Media Usage	
Never	15 (5%)
Monthly	21 (7%)
Weekly	49 (17%)
Daily	174 (58%)
Hourly	37 (13%)
Social Media Discussion Activity	
Not very active	39 (13%)
Not active	62 (21%)
Neither active nor inactive	54 (18%)
Active	117 (40%)
Very active	24 (8%)

New Table 4.5 (continued)

Online Shopping Frequency	
Never	1 (0%)
Rarely	15 (5%)
Sometimes	85 (29%)
Often	150 (51%)
Always	45 (15%)

ANOVA

The 236 participants exposed to a brand response message (Analytical, Clout, Authenticity, Negative Emotional Tone) were used in the analysis. Since consumers may receive numerous message types from brands, it is important to know which messages are able to resonate and which messages do not have their intended impact. The ANOVA assessed the most effective message type in increasing perceived genuineness – *the view that the interaction seems natural* – (Zhu and Meyers-Levy, 2009) and the realism of the brand response. Genuineness was chosen to measure the participants' view that a brand's response appears to be coming from a real person. The results of the ANOVA showed a statistically significant difference among the four LIWC treatments on both the perceived genuineness of the brand response ($F(3, 236) = 4.62, p < .01$) and realism ($F(3, 236) = 5.212, p < .01$).

The ANOVA and Least Significant Difference (LSD) post hoc test indicated that the Emotional Tone response ($M = 3.87$) was significantly less genuine than both the Clout response ($M = 4.95, p < .01$) and the Authentic response ($M = 4.90, p < .01$) while not significantly different from the Analytical response ($M = 4.31, p = .19$). Additionally, respondents found the solely negative response from the brand to be significantly less realistic ($M = 3.52$) than the Analytical ($M = 4.67$), Clout ($M = 4.27$), and Authentic ($M = 4.51$) treatments. Therefore, messages that use Clout or conveyed Authenticity exhibited greater sincerity and genuineness by consumer

vigilantes, whereas messages that only convey a negative Emotional Tone alone were seen as less suitable. Overly Analytical messages are not seen as genuine. Moreover, using social terms such as (I, we, us, you) conveys status (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2013) while also creating a sense of community during a crisis event (Baker and Hill, 2013). Since the Clout response builds connections with the receiver, it can be used to encourage collaboration with the vigilante (Wolfgang, 2018). Therefore, the social Clout treatment is used in Study 2 to provide a realistic and genuine response to participants.

Procedure

Study 2 employs an experimental research methodology using a survey to test the hypothesized relationships. The researcher recruited 112 respondents from an online panel on Amazon's M-Turk. To ensure quality online respondents, various attention checks, respondent screening policies, and a requirement of written text were required. Also, only U.S. respondents with a 95% work acceptance rate were considered for participation, and a nominal wage was paid to reduce misrepresentation (Hulland and Miller, 2018; Sharpe Wessling, Huber, and Netzer, 2017). The manipulation checks showed that respondents were aware when the brand responded to them, and when the brand did not, by stating, "The company responded to me" ($t = 12.26, p < .01$) and "The company sent me a message" ($t = 18.27, p < .01$). Therefore, the manipulation checks are successful.

Participants were asked to answer survey questions regarding the believability of the message ($\alpha = .92$; Saucier and Webster, 2010), justification in sharing the message ($\alpha = .96$, Inman and Zeelenberg, 2002), and likelihood to share misinformation ($\alpha = .94$, Daunt and Harris, 2011). A confirmatory factor analysis was then performed using AMOS 26 to assess the unidimensionality, convergent validity, and discriminant validity of the latent constructs before

creating composite measures. The results of the analysis indicated an acceptable fit (Marsh, Hau, and Wen, 2004) of the model to the data ($\chi^2=111.07$, $df = 41$, $CFI = .98$, $TLI = .97$ $IFI =.98$, $RMSEA = .07$). Table 4.6 shows a complete list of results from the CFA, and Cronbach's α for each construct from the ANOVA.

Table 4.6 CFA Results, t-values, and Cronbach's α

Scales	α	Standardized Factor Loadings	t-values
Justifiability of the decision - (adapted from Inman and Zeelenberg, 2002)	0.96		
Strongly Justifiable		0.90	31.23
Very Logical		0.84	24.22
Acceptable		0.96	43.94
Reasonable		0.97	*
Vigilantism - (adapted from Saucier and Webster, 2010)	0.92		
I feel as if it is my duty to enlighten other people.		0.91	*
I feel that my ideas should be used to educate others.		0.88	21.59
I feel a social obligation to voice my opinion.		0.88	21.33
Sharing misinformation (misbehavior intentions) - (adapted from Daunt and Harris, 2011)	0.94		
I would spread the information even though there is no genuine problem.		0.91	*
I would share exaggerated information when discussing the company.		0.88	23.10
I would share the unverified information to other consumers		0.89	24.14
I would share information in a way that the company may find unacceptable		0.88	23.46

Model fit statistics: $\chi^2 = 111.07$, $df = 41$, $p < 0.01$, $\chi^2/df = 2.71$, incremental fit index (IFI) = 0.97, Tucker-Lewis Index (TLI) = 0.98, comparative fit index (CFI) = 0.97, root mean square error of approximation (RMSEA) = 0.07.

All factor loadings have a p-value < 0.01

* denotes a constrained relationship to 1.00 for identification

Following Fornell and Larcker's (1981) criterion, all constructs displayed discriminant validity as the average variance extracted for each construct was higher than the squared correlations between any pairs of constructs (see Table 4.7). A latent common method factor test did not suggest any meaningful presence common method variance ($\Delta\chi^2/df=1 < 4$) (Podsakoff, MacKenzie, Lee, and Podsakoff, 2003).

Table 4.7 Means, Standard Deviations, AVEs, and Shared Variance:

	Mean	SD	1	2	3
Justifiability of the Decision	5.06	1.23	0.85		
Vigilantism	5.25	1.47	0.03	0.79	
Sharing Misinformation	4.24	1.81	0.08	0.30	0.79

Note: Average Variance Extracted is listed in the diagonals of the table

Note: AVEs including all treatments

Results

A simple regression shows the level of consumer vigilantism (0.19, $t = 4.04$, $p < 0.01$) and justification in their decision (.44, $t = 10.65$, $p < 0.01$) both significantly influence the likelihood to share self-created misleading information with others. Consumers who act more vigilantly online and who justify their actions in sharing misinformation are more likely to share self-created misleading information about a brand. The analysis uses PROCESS Model 3 to capture the moderated moderation effect.

*Vigilantism*Justification*

The interaction effect for Vigilantism*Justification is significant (0.10, $t = 2.26$, $p < 0.05$), supporting Hypothesis 4. However, the direct effects of vigilantism (-0.20, $t = -1.02$, $p = .31$) and justification (0.20, $t = 1.02$, $p = 0.32$) are non-significant. This finding suggests that the interaction becomes increasingly important in driving the likelihood to share self-created misleading information. The vigilante must justify their actions to share the misinformation for autoinoculation to occur.

Implementing the Quarantine

*Vigilantism*Justification* Brand Response*

The three-way interaction between Vigilantism*Justification*Response is significant (-0.12 , $t = -3.37$, $p < .01$). Additionally, the test of the highest order unconditional effects is also significant ($\Delta R^2 = 0.04$, $F = 11.36$, $[1, 104]$, $p < .01$) and confirms the presence of moderated moderation supporting Hypothesis 5. Therefore, the response from the brand becomes important in reducing the autoinoculation and likelihood that misleading information is shared. The conditional interactions ($X*W$) at values of Z were significant within the “non-response” treatment (0.10 , $F [1, 104] = 5.11$, $p < .05$) and within the “Clout response” treatment ($-.15$, $F [1, 104] = 6.31$, $p < .05$). This finding suggests that those who did not receive a message from the company were more likely to feel justified in creating and sharing misleading information.

Conversely, those who received the Clout response from the brand were less likely to feel justified in creating and sharing misleading information. This effect is especially apparent when a vigilante scores high on justification and does not receive any response from the brand (0.42 , $t = 2.97$, $p < 0.01$ CI Low = $.14$, CI High = 0.69). However, when a vigilante has a high level of justification and is met with an effective preemptive brand response, this effect can be reversed (-0.44 , $t = -2.27$, $p < .05$, CI Low = -0.82 , CI High = -0.05). The floodlight analysis provided a Johnson-Neyman point (Spiller et al., 2013; Johnson and Neyman, 1936). A Johnson-Neyman point is found when justification is greater than 3.73 on the 7-point scale ($p < .05$). This finding suggests that when justification is 3.73 or greater, the interaction effect between vigilantism and the treatment becomes negative.

When an individual is left unchecked, they can justify their actions and then more effectively autoinoculate against external information. This autoinoculation occurs because the

vigilante is able to incubate and let the viewpoint strengthen. When the viewpoint solidifies, they may then share misinformation and easily defend their actions to others. However, when the brand responds effectively and preemptively, the autoinoculation is unable to develop as the vigilante is faced with new contradictory information. Since the vigilante may have difficulty justifying their motives, they are less likely to share self-created misinformation with others. When the response is strong enough, they may actively contradict other statements they may have made. The results of Study 2 are available in Table 4.8.

Table 4.8 Study 2 Results

Hypothesized Relationship	Unstandardized Estimates	t-Values					Hypothesis Supported?
H4: Vigilantism*Justify=> Share Message	0.10	2.25*					Yes
Vigilantism => Share Message	-0.20	-1.02					
Justify => Share Message	0.20	1.02					
H5: Vigilantism*Justify*Response=> Share Message	-0.12	-3.37*					Yes
Treatment	Effect	se	T	p	LLCI	ULCI	
0	0.42	0.14	2.97	0.004	0.14	0.69	
1	-0.44	0.19	-2.28	0.03	-0.82	-0.06	

Discussion

These results build upon the findings of Study 1 by indicating that the source of the inoculation can be contained. A vigilante can have a higher propensity to share misinformation with others, and higher levels of justification strengthen this likelihood. However, the brands' providing a prompt response that attempts to create an ally out of a potential enemy can reduce the possibility that the necessary justification for autoinoculation is formed. If the brand provides overly negative messages or "matches" a belligerent individual's tone, this strategy can backfire.

Additionally, brand responses that are not deemed genuine do not significantly reduce the justification for the vigilante's actions.

If the individual is challenged early enough, before their views are able to be fully justified, then they will be less likely to share that misinformation because they are unable to justify this action to themselves. Therefore, online consumers may be more hesitant to share misinformation early when they feel that the brand actively conveys and manages expectations to other consumers. Therefore, brands attempting to communicate good faith intentions can communicate directly to both vigilantes and the broader community.

General Discussion and Implications

By exploring the role of brand quarantine responses in attenuating the intentions to share misleading information, a better understanding of the cognitive processes of autoinoculation is developed. This understanding contributes to the development of inoculation theory by providing evidence that early autoinoculation can be fleeting. These two studies provide evidence that the vigilante can play a role in preventing the spread of inoculation to others. Moreover, vigilantes must be overt in gaining the attention of a community in offsetting the effect of the original message. Although more discreet responses are likely more common, they do not have the same level of impact in halting a Negative Cascade. The thought listings that were included at the end of each survey in the individual studies provided additional insight. For instance, one participant that did not receive a rebuttal statement stated that:

“ Because I am aware of the amount of false reviews and negative comments posted by competitors, I would usually take a negative comment about a company with a grain of salt.”

“I think customer reviews can be manipulated and fraudulently misleading to manipulate customer behaviors.”

Other participants, in the discreet treatment, stated that it is difficult to repair a first impression and that once a message is seen by a large group, it may be difficult to offset. Moreover, participants believe that this is a problem that can quickly develop online due to malicious consumers. Additionally, another participant stated,

“There are bored people out there who get a kick out of getting other people riled up which I consider a mob effect even if none of this is true. There should always be a moderator of reviews to protect the company's brand”.

Participants also found parallels between misinformation-sharing about brands and the sharing of misleading news stories.

“This is similar to what is happening in the spreading of mainstream news. Stories, whether true or false, are being believed right at the start. No one waits to see how it pans out. And they definitely don't do their own independent research.”

Although the more overt retraction helped to attenuate the Negative Cascade, it also can present challenges for brands to initiate. For instance, one participant said the following:

“I didn't believe the scenario was believable because there's no way someone on the internet would admit a mistake like that. They'd either delete their post or their reddit [sic] account.”

Yet, this scenario described an actual occurrence from January 2016. This insight provides evidence of the unique challenges to brands in attempting to persuade consumer responses after a cascade has taken effect. Once a post is created, vigilantes may be unlikely to contradict themselves. However, building a pseudo-ally – *party that reluctantly joins ranks with*

the brand – and providing adequate support to the vigilante can provide the outlets to offset misinformation. This finding suggests that those who have begun to inoculate others can break the effect if they are able to un-inoculate themselves.

Study 2 showed a brand should not attempt to match the negative affect of an online consumer when trying to dissuade misleading information sharing. Providing an overly negative response or overly Analytical response is seen as less genuine and can make matters worse. For instance, one participant stated:

“If a company is less friendly [sic] then people are more likely to share negative information.”

Alternatively, brands that communicate diligently and effectively are deemed more genuine and can be seen in a more favorable light.

“When a company takes time to have a live person address a situation I find I am more likely to do future business with them rather than companies that use automated responses.”

Theoretical Implications

These findings extend the inoculation literature by showing the role of the individual in reversing the consumer herd inoculation and the role of the brand in quarantining the individual long enough to prevent autoinoculation. Moreover, these findings extend inoculation theory as a dual-process model of decision-making. Since dual-process models rely on automatic and analytic components, the autoinoculation, and herd inoculation, requires both as well. A vigilante may first feel a desire to share a misleading message, whether intentional or not. Although a snap judgment through automatic processes can drive the sharing of a message, analytical decision-making is required to justify the message to themselves and others. When the vigilante defends

the message, and it solidifies the intent of the automatic internal processes, they can cement their view.

The dual-process model of inoculation is expanded in this research to show that cognitive justification can be attenuated by a responding brand. A brand response during the vigilante's incubation stage (Banas and Rains, 2010), when the vigilante is deciding the merits of information sharing, can reduce the justification and attenuate the initial and future sharing of a misleading message since doing so can bring scrutiny from others (Kahneman and Frederick, 2002). For example, analytical resources require deliberate thought and rationalization for actions to be taken (Kahneman, 2011), and additional effort may be required to make moral judgments (Haidt, 2001). The findings here show that social Clout messages from a brand can facilitate a pseudo-ally relationship by conveying leadership, social status, and confidence (Pennebaker, 2019). These effective brand responses can prevent the vigilante from inoculating themselves and, thus, the larger group. The vigilante, acting in a pseudo-ally role, may then refrain from sharing a misleading message or attempt to correct a previously shared message.

Since dual-process inoculation requires cognitive justification for future sharing, a vigilante may act to challenge the prevailing content by overtly communicating to a group and correcting misinformation that may have been shared. This vigilante response may challenge group/community communication norms (Algesheimer, Dholakia, and Herrmann, 2005) and require more focus and control to be directed to updates in the discussion (Schau, Muñiz, and Arnould, 2009). However, persuading a vigilante to change a view can be incredibly challenging to initiate when direct admission of fault is required (Rosenbaum et al., 2011). It is crucial at this juncture that a brand provides the necessary support to the vigilante and pursue collaborative interactions. Once other consumers are forced to direct additional cognitive resources to

challenge the original versus new message from the vigilante, they may determine that future sharing is not justified. This, in effect, may break the consumer herd inoculation. Therefore, brands need ambassadors to show goodwill while also swaying those that may otherwise cause harm. Brands need allies, and sometimes these allies may come from unlikely places.

Managerial Implications

These findings provide implications for managers who aim to understand misinformation sharing and the response required to quarantine misinformation. The real damage from a lone vigilante comes from more substantial group sharing. Once a broader community is exposed to a misleading message, and corresponding messages begin to reinforce that message, a Negative Cascade reinforces both the initial belief in the message itself and justification in spreading the message by other consumers. As consumers spread the misinformation, the herd inoculation and shared values can be further reinforced and create an echo chamber of misinformation. A corrective vigilante message that is too discreet may provide less guidance in halting the future sharing of a message and act as ignorable background noise in the online discussion (Phan and Godes, 2018).

The best route for a brand to take in quarantining the vigilante is to provide a message that profoundly conveys social Clout and attempts to build a pseudo-ally. This message should include social words such as we and us, convey confidence and knowledge in the discussion, and provide an apologetic tone if a mistake was made. By addressing the individual early, the effect of justification in driving autoinoculation is attenuated, and misinformation sharing can be mitigated. Essentially, the vigilante, unable to offset the potential views of others in the presence of the message, may direct efforts to act as a benefactor to correct a previous misstatement and act as a watchdog for the future misinformation issues (Larson and Denton, 2014). Thus, a brand

should provide support to the vigilante to help them build up the necessary overt response required to attenuate a Negative Cascade.

Community participants may be influenced by overt gestures from those who are correcting a misleading post, but this action is not always expected. Once an individual posts a message that causes potential harm, they will most likely not contradict themselves. Convincing consumers to make these grand gestures can be nearly impossible. One solution to this issue could be for brands to provide incentives or share information with the community about the vigilante in a positive light to reinforce “good” behavior.

Brands should diligently monitor online communities and the conversations that occur. If resentment is found and issues arise, brands can initiate a preemptive response to a potentially misleading vigilante to reduce the likelihood that a message is shared. A “Keeping your enemies closer” mindset can be useful in creating allies through social Clout and conveying a sense of “we-ness.” For example, the Clout message treatment, “*Hey there, I'm afraid that the price we quoted was incorrect. Can you please cancel your order?*” deftly used social words such as “I, we, and you” and helped to reduce the likelihood of misleading information being shared. Since a malicious action directed toward a liked brand is more difficult to justify, vigilantes may opt to search for new information before initiating a misleading message. The Clout brand response hinders true vigilantes from initiating a response and reduces the strength of a Negative Cascade.

While this approach may not lead to positive information sharing, it may halt misinformed vigilantes from spreading a misleading message and put a spotlight on true trolls. For instance, those with solely malicious intent may not require the same justification in generating a misleading message and might be more noticeable to the rest of the community. When others are exposed to both a brand and vigilante’s response, they gain additional evidence

to make informational assessments. Additionally, when the vigilante becomes aware that others are watching, and when responses can be compared, they may feel that they can be more easily caught in a lie. A study participant stated the following:

“On social media accounts, I see a lot of customers complaining about every little thing and trying to make the company look bad. I do like to see how the company personally responds to everyone though.”

Whether a complaint is legitimate, a claim is factual, or experience is real, a brand must manage online information-sharing. An active online presence can disarm harmful intentions and negative affect directed toward a brand. When countering information is made available quickly, autoinoculation, and as a result, herd inoculation has difficulty gaining momentum. In either scenario, brands should monitor the information sharing that takes place online, determine the accuracy of these statements, participate in discussions, and respond quickly when deviations are found.

Limitations and Future Research

While this research provides important implications for both practitioners and academicians, limitations and opportunities for future research are present. The scenario focused on a brand accused of manipulating online reviews. Future studies should focus on further brand transgressions across product categories. Additionally, the response treatments were created using LIWC and focused on creating a message high in only one LIWC summary variable (Analytical, Clout, Authenticity, or Negative Emotional Tone). Since messages between brands and consumers can be quite complex and situation-specific, future studies could focus on combinations of message types and consumer situations. In doing so, more nuanced comparisons can be made into what constitutes

an appropriate brand response, additional abnormalities in communication, and the elimination of misleading information.

Future research should focus on the types of consumers that share misleading information online. For instance, consumers with high propensities for the Dark Triad (Machiavellianism, Narcissism, and Psychopathy) (Jonason and Webster, 2010), may not respond to the same mechanisms of justification and the associated brand quarantine as other vigilantes. These consumers may require additional resources and challenges to address and may require a much more collaborative community response to extinguish. Future research should focus on the use of vigilantes as pseudo-allies to help address overly malicious and damaging consumer behavior and the role of brand enthusiasts as both moderators and advocates of productive consumer discussions.

Additionally, the treatments used a fictitious brand created by the author to prevent bias from affecting the results of the studies. Future research could focus on established brands in determining the ability of an individual to initiate a Negative Cascade and the effort required for the established brand to halt it. Additionally, the treatment in Study 2 focuses on the potential exploitation of a pricing mistake and the initiator's likelihood to share self-created misleading information. Other scenarios and situations should be tested to determine other driving factors of misinformation, such as brand hatred, self-amusement, contempt, or service failures. Additionally, future studies can implement facial recognition and eye-tracking to biometrically triangulate the negative/positive responses that emerge.

M-Turk samples were used in both Studies 1 and 2. Since the study focused on general online shopping situations, the sampling method was deemed appropriate as a wide range of consumers participate in both shopping situations. Also, the author set parameters that would only include U.S. respondents, and only those with a 95% work acceptance rate were considered for

participation. While care was taken to provide a scenario appropriate for these samples, and the methods employed best practices for M-Turk (Hulland and Miller, 2018; Hulland, Baumgartner, and Smith, 2018; Sharpe Wessling, Huber, and Netzer, 2017), future research might replicate these findings using a different sampling approach.

Conclusion

By understanding the role of a vigilante in influencing a cascade of information, the findings provide an understanding of the direct communication that can offset misinformation with the potential to develop into a brand crisis. This understanding also contributes to knowledge by exploring the most effective ways to address these individuals. The studies provide additional insights for companies by uncovering the role of a vigilante, albeit a negligent one in the initial misinformation-sharing, in reversing online discourse. The findings provide insights into the responses that can reverse autoinoculation and thus halt consumer herd inoculation.

CHAPTER V

OVERVIEW AND CONCLUSION

While social media and consumer interaction can drive brand value and success, it must be monitored carefully to minimize the impact of malicious and misinformed individuals. Consumer-inspired negative information is disconcerting for over 75% of business owners who state that online reviews are vital, and one-sixth of those owners said bad reviews could potentially ruin their business (Pickard-Whitehead, 2017). Whether it be post-truth politics (Roberts, 2016), truthiness in information sharing (Meddaugh, 2010; Munger, 2008), harmful information-sharing (Newman, Garry, Bernstein, Kantner, and Lindsay, 2012), or overt trolling behavior (Buckels et al. 2014), brands have a number of challenges in managing online consumer discourse. In isolated incidents, misinformation may have little effect, but deviant behaviors can rely on large groups to “bomb” average review scores on Yelp, TripAdvisor, Amazon, and IMDB to cause intentional harm to brands (Gonimah, 2019). Moreover, these consumers can also develop anti-brand communities (Hollenbeck and Zinkhan, 2006), anti-company websites (Kucuk, 2010), and serve as a match for a firestorm of negative WOM (Herhausen et al. 2019).

Consumers who use online platforms to create and share misleading information to meet an end are able to autoinoculate themselves from contradictory brand messages and then begin the herd inoculation of others. When this occurs, a brand must come to its own defense regardless of the inaccuracy or misleading nature of the original post. Therefore, brands must be

diligent in proactively addressing these individuals. Through the lens of inoculation theory, decision aids are provided to brands through a more robust understanding of the autoinoculation of the individual, the herd inoculation of the online group, and the potential remedy provided by a brand quarantine.

Answering the surge in interest in dealing with the more dysfunctional aspects of consumer behavior (Payne and Frow, 2017), this dissertation aimed to address the dark side of consumer behavior and aid in diagnosing issues that brands face online. This dissertation, consisting of three separate essays, represents an attempt to aid brands by addressing how misleading information is created, how it is disseminated, and how it can be eliminated. Individually, each essay provides a building block that explains the source of misinformation, how bystanders may be swayed, and the numerous strategies brands have to alleviate the severity of misleading information sharing.

Essay One: Autoinoculation of Poster Zero: The Role of the Dark Triad and Vigilantism in the Creation of Misleading Information

Essay One employed qualitative and quantitative methods to extend the lens of inoculation theory (McGuire, 1968) by exploring autoinoculation as both automatic and analytical cognitive processes within a dual-process model of decision-making (c.f. Kahneman, 2011). The critical link between the Dark Triad (Machiavellianism, Narcissism, and Psychopathy) (Jonason and Webster, 2010; Jones and Paulhus, 2014) and misleading information sharing is established, the importance of both consumer proactivity and vigilantism (Saucier and Webster, 2010) is uncovered, and self-justification (autoinoculation) of behaviors is developed to understand the reasons why consumers engage in malicious behavior that may

cause a fauxstorm –*a sudden discharge of misinformation that gains credibility as a Negative Cascade is created and spreads the message to larger consumer groups.*

The qualitative netnographic content analysis, examining ten Reddit forums, explored the thoughts and feelings associated with malicious online behavior and the types of opportunities that these consumers pursue. The emerging themes were then tested using structural equation modeling. The findings suggest that consumers' Dark Triad propensities affect both proactivity and vigilantism and thus strongly influence the likelihood to create and then share purposely misleading information with other consumers when the individual feels justified in doing so.

The quantitative study extended research on the Dark Triad by showing Machiavellianism, Narcissism, and Psychopathy, together, create a formative Dark Triad construct. Additionally, the formative Dark Triad construct is a predictor of consumer proactivity in seeking opportunities and utilizing online platforms to facilitate vigilantism. The propensities, when paired with cognitive justification for their actions, influence the likelihood of a consumer to create and then share misleading information with others. These findings suggest that even individuals with high levels of Dark Triad traits still justify their actions to themselves and inoculate themselves from contradictory views. Therefore, although the Dark Triad does influence the malicious intent of consumers, proactivity and vigilantism can further explain this relationship. Moreover, this effect is strengthened when the poster can justify their actions.

Essay Two: The Herd and the Bystander: The Role of Availability Cascades in the Diffusion of Misleading Information in Online Environments

Essay Two explored the phenomenon of availability cascades in inducing groups of consumers to both accept and share misleading messages. As information becomes more available to others online, the original poster can expend less effort for the message to spread,

thus causing a fauxstorm of misinformation. This factor is essential for brands to understand since an active brand community may be able to take advantage of the same phenomenon to mitigate harmful information sharing. However, not all brands have active brand communities, and consumer bystanders may only see negative information that confirms the original poster's point.

Inoculation theory as a dual-process model was used as a theoretical lens to determine which message types, (Analytical, Clout, Authentic, or Emotional Tone), have the highest likelihood of acceptance by other consumers. The findings suggest that a message high in authenticity when conveying a negative tone is most believable by others. Additionally, the presence of a confirmation/disconfirmation cascade acted to strengthen/weaken the corresponding believability and thus likelihood of sharing the message with others. The good news for brands is that a cascade works for both the benefit of the poster and the benefit of the brand. But, challenges in creating counter messages are present.

The bystander effect suggests that individuals may have difficulty acting in situations in which others are present. This effect was on display when participants were exposed to information that contradicts the original poster. For instance, when consumers are exposed to a negative message that is confirmed by a larger group, a truth-default effect enhances the likelihood to post an additional negative message and greatly reduces the likelihood to post anything positive. Moreover, those exposed to matching negative treatments felt that they would be more likely to join in the discussion. However, although a Positive Cascade that contradicts the original poster increased the likelihood to post a positive message, the effect was weaker than those in the Negative Cascade. This finding suggests that conflicting information makes decisionmaking more difficult and can drive the bystander effect in those that would like to post

a positive message for the brand. Essay Two provides evidence that information cascades are crucial to hindering misinformation sharing yet are complicated to initiate. The cascades drive continual information sharing, but driving continual participation to completely attenuate the herd inoculation requires additional actions from the brand.

Essay Three: Quarantining the Vigilante: The Importance of Early Brand Response in Eliminating the Effects of Misleading Information

With the importance of the individual consumer established in Essay One, and the role of the group found in Essay Two, Essay Three examined how halting a vigilante's actions before the cascade has time to take effect can greatly reduce the chances of both autoinoculation and herd inoculation. This essay examined the procedures that, if carried out early enough, can influence consumers before the critical mass of the availability cascade takes effect. Essay Three examines those exhibiting vigilante behaviors and propensities who may be misguided in their attempts to influence others through misinformation. A quarantine is introduced as a possible remedy for the misleading information by isolating the vigilante's message before the Negative Cascade spreads the inoculation throughout the community.

The quantitative experimental analyses provided insights into the power that a vigilante can have in offsetting a Negative Cascade once it starts. Although a discreet corrective message within the forum has a negative interaction effect with information-sharing, it was nevertheless non-significant. However, an overt corrective message from the vigilante poster significantly reduced the likelihood of others in sharing the vigilante's negative message and helped to offset the Negative Cascade. However, this requires the original poster to admit fault, and this admission of wrongdoing may not easily occur. Therefore, brands must act diligently to not only

influence posters to correct information but also build connections and alliances with these would-be adversaries.

Initiating a quarantine early can reduce the poster's necessary cognitive justification for the negative actions and thus prevent autoinoculation from occurring. Since inoculation requires an incubation window for views to be cemented, a brand can act to prevent its full effect. The study showed that those with higher levels of vigilantism were likely to create and share misleading information, and justification (autoinoculation) helped to drive this effect. By introducing a quarantine treatment through a preemptive response to the poster from the brand, the three-way interaction effect becomes negative and reduces the likelihood that misinformation is shared. The study found that messages conveying social words and clout had the greatest effect in attenuating the autoinoculation. By responding proactively to the vigilante and creating a pseudo-ally, the brand can act to halt autoinoculation at the source and, in turn, halt the herd inoculation.

Synthesis and Final Remarks

The three essays in this dissertation represent an effort to understand and explain the mechanisms that drive the creation, dissemination, and elimination of misleading information that is propagated online. This effort extends inoculation theory with both qualitative and quantitative techniques that help to articulate the intent of individuals, their influence on others, and the various responses from both brands and vigilantes. The fact that misinformation spreads is less daunting than the threat that its spread cannot be halted. If a consumer spreads a rumor and then finds information that contradicts it, causing them to change their view, then the contagion was not as daunting. However, if the misinformation is spread and then

counterinformation still does not reduce this belief, real problems can arise. This dissertation aims to understand these processes and provide solutions to brands.

By pulling from real discussions online, Essay One determines that individuals that exhibit the Dark Triad propensities are proactive in seeking out ways to benefit themselves and act as vigilantes when sharing their beliefs. The essay provides insights into the antecedents and outcomes of these types of propensities. Although higher levels on the Dark Triad enhanced this relationship, all consumers, when shielded by anonymity, can act more maliciously than they otherwise would in real-life. Thus, an online disinhibition effect makes this a uniquely online problem for brands to address. When the poster autoinoculates by justifying their actions in the face of potential scrutiny, then the first message can be the spark to initiate a Negative Cascade and a corresponding fauxstorm.

Essay Two continues the contribution of Essay One by understanding how the misinformation spreads online. Bystanders exposed to the information do not necessarily know that the information is intentionally misleading, so the original poster can mask intent by appearing authentic and humble. When others appear to confirm the message, a would-be bystander assumes a truth default effect and then may choose to share the message with others and thus continue the spread of the herd inoculation. However, a Positive Cascade may attenuate the effect of the herd inoculation, but this effect is strongest when the original poster seems less authentic. The problem that arises for brands is that conflicting information requires more analysis and cognitive resources, and bystanders may be less likely to generate the corresponding positive information to offset the herd inoculation.

Essay Three introduces treatments to both halt the cascade and generate a necessary quarantine around the vigilante poster to prevent the initial message from being spread. Building

a pseudo-ally relationship helps to prevent the brand from becoming a target and can also influence other posts and corrections that can halt herd inoculation caused by the Negative Cascade. Brands must be diligent and enact both automated and more manual means for providing remedies for misinformation online.

Overall, the essays contained in this dissertation provide both theoretical and empirical lenses for future studies in malicious consumer behavior, inoculation, and how misinformation spreads online. The total culmination of research presented is intended to be a starting point to understanding how misinformation can be addressed and contained. By continuing this research stream, future development can provide essential insights to consumers, practitioners, and researchers.

REFERENCES

- Algesheimer, R., Dholakia, U. M., & Herrmann, A. (2005). The social influence of brand community: Evidence from European car clubs. *Journal of Marketing*, 69(3), 19-34.
- Akpınar, E., Verlegh, P. W., & Smidts, A. (2018). Sharing product harm information: The effects of self-construal and self-relevance. *International Journal of Research in Marketing*, 35(2), 319-335.
- Alfano, M. (2011). Explaining away intuitions about traits: Why virtue ethics seems plausible (even if it isn't). *Review of Philosophy and Psychology*, 2(1), 121-136.
- Al-Khatib, J. A., Vitell, S. J., & Rawwas, M. Y. (1997). Consumer ethics: A cross-cultural investigation. *European Journal of Marketing*, 31(11/12), 750-767.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411.
- Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2019). The future of social media in marketing. *Journal of the Academy of Marketing Science*, 1-17.
- Baker, S. M., & Hill, R. P. (2013). A community psychology of object meanings: Identity negotiation during disaster recovery. *Journal of Consumer Psychology*, 23(3), 275-287.
- Banas, John A; Rains, Stephen A (2010). "A Meta-Analysis of Research on Inoculation Theory". *Communication Monographs*. 77 (3): 281–311.
- Barnett, Michael. (2014). Warning: how social media is a danger to brands (Retrieved from (<https://www.marketingweek.com/warning-how-social-media-is-a-danger-to-brands/>))
- Baron, R. A. (1998). Cognitive mechanisms in entrepreneurship: Why and when entrepreneurs think differently than other people. *Journal of Business venturing*, 13(4), 275-294.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173.
- Barasch, A., & Berger, J. (2014). Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research*, 51(3), 286-299.

- Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. *Handbook of Social Cognition, 1*, 1-40.
- Barr, W. B. (2013). An evidence based approach to sports concussion: confronting the availability cascade.
- Bateman, T. S., & Crant, J. M. (1993). The proactive component of organizational behavior: A measure and correlates. *Journal of Organizational Behavior, 14*(2), 103-118.
- Bayl-Smith, P. H., & Griffin, B. (2014). Age discrimination in the workplace: identifying as a late-career worker and its relationship with engagement and intended retirement age. *Journal of Applied Social Psychology, 44*(9), 588-599.
- Bechwati, N. N., & Morrin, M. (2003). Outraged consumers: Getting even at the expense of getting a good deal. *Journal of Consumer Psychology, 13*(4), 440-453.
- Bee, C. C., & Madrigal, R. (2013). Consumer uncertainty: The influence of anticipatory emotions on ambivalence, attitudes, and intentions. *Journal of Consumer Behaviour, 12*(5), 370-381.
- Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of Marketing Research, 49*(2), 192-205.
- Bergstrom, K. (2011). "Don't feed the troll": Shutting down debate about community expectations on Reddit.com. *First Monday, 16*(8).
- Berthon, P., Pitt, L., & DesAutels, P. (2011). Unveiling videos: Consumer-generated ads as qualitative inquiry. *Psychology & Marketing, 28*(10), 1044-1060.
- Bernritter, S. F., van Ooijen, I., & Müller, B. C. (2017). Self-persuasion as marketing technique: the role of consumers' involvement. *European Journal of Marketing, 51*(5/6), 1075-1090.
- Beukeboom, C. J., Kerkhof, P., & de Vries, M. (2015). Does a virtual like cause actual liking? How following a brand's Facebook updates enhances brand evaluations and purchase intention. *Journal of Interactive Marketing, 32*, 26-36.
- Binns, A. (2012). DON'T FEED THE TROLLS! Managing troublemakers in magazines' online communities. *Journalism Practice, 6*(4), 547-562.
- Bjorklund, T., Bhatli, D., & Laakso, M. (2013). Understanding idea advancement efforts in innovation through proactive behavior. *Journal of Research in Marketing and Entrepreneurship, 15*(2), 124-142.
- Blodgett, J. G., Wakefield, K. L., & Barnes, J. H. (1995). The effects of customer service on consumer complaining behavior. *Journal of services Marketing, 9*(4), 31-42.

- Boyd III, H. C., & Helms, J. E. (2005). Consumer entitlement theory and measurement. *Psychology & Marketing, 22*(3), 271-286.
- Breitsohl, J., Roschk, H., & Feyertag, C. (2018). Consumer Brand Bullying Behaviour in Online Communities of Service Firms. In *Service Business Development* (pp. 289-312). Springer Gabler, Wiesbaden.
- Brüggen, E. C., Foubert, B., & Gremler, D. D. (2011). Extreme makeover: Short-and long-term effects of a remodeled servicescape. *Journal of marketing, 75*(5), 71-87.
- Brown, S., Kozinets, R. V., & Sherry Jr, J. F. (2003). Teaching old brands new tricks: Retro branding and the revival of brand meaning. *Journal of Marketing, 67*(3), 19-33.
- Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences, 67*, 97-102.
- Calhoun, R. P. (1969). Niccolo Machiavelli and the twentieth century administrator. *Academy of management Journal, 12*(2), 205-212.
- Carrigan, M., & Attalla, A. (2001). The myth of the ethical consumer—do ethics matter in purchase behaviour?. *Journal of consumer marketing, 18*(7), 560-578.
- Chalkiti, K., & Sigala, M. (2008). Information sharing and knowledge creation in online forums: The case of the Greek online forum 'DIALOGOI'. *Current Issues in Tourism, 11*(5), 381-406.
- Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research, 50*(4), 463-476.
- Cherrier, H. (2009). Anti-consumption discourses and consumer-resistant identities. *Journal of Business Research, 62*(2), 181-190.
- Chester, D. S., & DeWall, C. N. (2017). Combating the sting of rejection with the pleasure of revenge: A new look at how emotion shapes aggression. *Journal of Personality and Social Psychology, 112*(3), 413.
- Cheung, G. W., & Lau, R. S. (2017). Accuracy of parameter estimates and confidence intervals in moderated mediation models: A comparison of regression and latent moderated structural equations. *Organizational Research Methods, 20*(4), 746-769.
- Choi, J., Li, Y. J., Rangan, P., Chatterjee, P., & Singh, S. N. (2014). The odd-ending price justification effect: the influence of price-endings on hedonic and utilitarian consumption. *Journal of the Academy of Marketing Science, 42*(5), 545-557.
- Christie, R., & Geis, F. (1970). Scale construction. *Studies in machiavellianism, 10*-34.

- Cialdini, R. B. (1984). *Influence: The psychology of persuasion*. New York. William Morrow and Company, Inc.
- Clark, T. (2019). *Social Media Trolls: A Practical Guide for Dealing With Impossible People*. Retrieved from (<https://blog.hootsuite.com/how-to-deal-with-trolls-on-social-media/>)
- Cohn, M.A., Mehl, M.R., & Pennebaker, J.W. (2004). Linguistic Markers of Psychological Change Surrounding September 11, 2001. *Psychological Science*, 15, 687-693.
- Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, 61(12), 1250-1262.
- Coles, B. A., & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior*, 60, 233-244.
- Compton J., & Ivanov, B. (2012). Untangling threat during inoculation-conferred resistance to influence. *Communication Reports*. 25, 1–13.
- Compton, J., & Pfau, M. (2009). Spreading inoculation: Inoculation, resistance to influence, and word-of-mouth communication. *Communication Theory*, 19(1), 9-28.
- Coombs, W. T. (2007). Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate reputation review*, 10(3), 163-176.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*.
- Cremer, J. (2018). Russia uses ‘information warfare’ to portray GMOs negatively. Retrieved from <https://allianceforscience.cornell.edu/blog/2018/02/russia-uses-information-warfare-portray-gmos-negatively/>
- Daugherty, T., & Hoffman, E. (2014). eWOM and the importance of capturing consumer attention within social media. *Journal of Marketing Communications*, 20(1-2), 82-102.
- Daunt, K. L., & Harris, L. C. (2011). Customers acting badly: Evidence from the hospitality industry. *Journal of Business Research*, 64(10), 1034-1042.
- Dawson, E.J., & Brashers, D. (1996). Information manipulation theory: A replication and assessment. *Communication Monographs*, 63(1), 70-83.
- DeAngelis, M., Bonezzi, A., Peluso, A. M., Rucker, D. D., & Costabile, M. (2012). On braggarts and gossips: A self-enhancement account of word-of-mouth generation and transmission. *Journal of Marketing Research*, 49(4), 551-563.

- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review, 14*(2), 238-257.
- Dekay, S. H. (2012). How large companies react to negative Facebook comments. *Corporate Communications: An International Journal, 17*(3), 289-299.
- Denegri-Knott, J. (2006). Consumers behaving badly: deviation or innovation? Power struggles on the web. *Journal of Consumer Behaviour, 5*(1), 82-94.
- Dhar, R., & Gorlin, M. (2013). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology, 23*(4), 528-542.
- Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*(2), 269-277.
- Dootson, P., Lings, I., Beatson, A., & Johnston, K. A. (2017). Deterring deviant consumer behaviour: when 'it's wrong, don't do it' doesn't work. *Journal of Marketing Management, 33*(15-16), 1355-1383.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of experimental social psychology, 33*(5), 510-540.
- Dubois, D., Rucker, D. D., & Tormala, Z. L. (2011). From rumors to facts, and facts to rumors: The role of certainty decay in consumer communications. *Journal of Marketing Research, 48*(6), 1020-1032.
- Dubois, D., Bonezzi, A., & De Angelis, M. (2016). Sharing with friends versus strangers: How interpersonal closeness influences word-of-mouth valence. *Journal of Marketing Research, 53*(5), 712-727.
- Edstrom, M. (2016). The trolls disappear in the light: Swedish experiences of mediated sexualised hate speech in the aftermath of Behring Breivik. *International Journal for Crime, Justice and Social Democracy, 5*(2), 96-106.
- Eisend, M. (2006). Two-sided advertising: A meta-analysis. *International Journal of Research in Marketing, 23*(2), 187-198.
- Elliott, R., & Jankel-Elliott, N. (2003). Using ethnography in strategic consumer research. *Qualitative market research: An international journal, 6*(4), 215-223.
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241.

- Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, 68(6), 1261-1270.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., & Kainbacher, M. (2011). The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137(4), 517.
- Fitzgerald, M. P., Russo Donovan, K., Kees, J., & Kozup, J. (2019). How confusion impacts product labeling perceptions. *Journal of Consumer Marketing*, 36(2), 306-316.
- Follett, P. A., & Neven, L. G. (2006). Current trends in quarantine entomology. *Annu. Rev. Entomol.*, 51, 359-385.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Fournier, S., & Lee, L. (2009). Getting brand communities right. *Harvard business review*, 87(4), 105-111.
- Friedmann, D. (2019). Correcting Information Asymmetry via Deep Consumer Information; Compelling Companies to Let the Sunshine In. *CONSUMER LAW & ECONOMICS*, Klaus Mathis & Avishalom Tor, eds., Springer (2019, Forthcoming).
- Fuhrmann, A. (2017). This Area Is [NOT] under Quarantine: Rethinking Southeast/Asia through Studies of the Cinema. In *Area Studies at the Crossroads* (pp. 251-268). Palgrave Macmillan, New York.
- Fullerton, R. A., & Punj, G. (1997). What is consumer misbehavior?. *ACR North American Advances*.
- Gallois, C., & Giles, H. (2015). Communication accommodation theory. *The international encyclopedia of language and social interaction*, 1-18.
- Ganesh, G. (1997). Spousal influence in consumer decisions: a study of cultural assimilation. *Journal of Consumer Marketing*, 14(2), 132-155.
- Garcia, S. M., Weaver, K., Darley, J. M., & Spence, B. T. (2009). Dual effects of implicit bystanders: Inhibiting vs. facilitating helping behavior. *Journal of Consumer Psychology*, 19(2), 215-224.
- Garcia, S. M., Weaver, K., Moskowitz, G. B., & Darley, J. M. (2002). Crowded minds: the implicit bystander effect. *Journal of personality and social psychology*, 83(4), 843.

- Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: employing structural equation modeling to test for construct validity. *Journal of Business Logistics*, 20(1), 33.
- Gatignon, H., & Robertson, T. S. (1985). A propositional inventory for new diffusion research. *Journal of consumer research*, 11(4), 849-867.
- Gaziel Yablowitz, M., & Raban, D. R. (2016). Investment decision paths in the information age: The effect of online journalism. *Journal of the Association for Information Science and Technology*, 67(6), 1417-1429.
- Golf-Papez, M., & Veer, E. (2017). Don't feed the trolling: rethinking how online trolling is being defined and combated. *Journal of Marketing Management*, 33(15-16), 1336-1354.
- Goodman, J. K., Cryder, C. E., and Cheema, A. (2013), "Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples," *Journal of Behavioral Decision Making*, Vol. 26 No. 3, pp. 213-224.
- Grégoire, Y., Legoux, R., Tripp, T. M., Radanielina-Hita, M. L., Joireman, J., & Rotman, J. D. (2018). What do online complainers want? An examination of the justice motivations and the moral implications of vigilante and reparation schemas. *Journal of Business Ethics*, 1-22.
- Grégoire, Y., Salle, A., & Tripp, T. M. (2015). Managing social media crises with your customers: The good, the bad, and the ugly. *Business Horizons*, 58(2), 173-182.
- Grau, S. L., & Folse, J. A. G. (2007). Cause-related marketing (CRM): The influence of donation proximity and message-framing cues on the less-involved consumer. *Journal of Advertising*, 36(4), 19-33.
- Gummesson, E. (2005). Qualitative research in marketing: Road-map for a wilderness of complexity and unpredictability. *European Journal of Marketing*, 39(3/4), 309-327.
- Gürhan-Canli, Z., & Maheswaran, D. (2000). Cultural variations in country of origin effects. *Journal of Marketing Research*, 37(3), 309-317.
- Guzmán, F., Paswan, A. K., & Van Steenburg, E. (2015). Self-referencing and political candidate brands: A congruency perspective. *Journal of Political Marketing*, 14(1-2), 175-199.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, 108(4), 814.
- Hair J. F., Black W. C., Babin, B.J. and Anderson R.E. (2010). *Multivariate Data Analysis* (7th edition), New Jersey: Pearson

- Hansen, N., Kupfer, A. K., & Hennig-Thurau, T. (2018). Brand crises in the digital age: The short-and long-term effects of social media firestorms on consumers and brands. *International Journal of Research in Marketing*, 35(4), 557-574.
- Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annu. Rev. Clin. Psychol.*, 4, 217-246.
- Harris, L. C., & Daunt, K. L. (2011). Deviant customer behaviour: A study of techniques of neutralisation. *Journal of Marketing Management*, 27(7-8), 834-853.
- Harrison, A., Summers, J., & Mennecke, B. (2016). The effects of the dark triad on unethical behavior. *Journal of Business Ethics*, 1-25.
- Hartnett, S. J. (2010). Communication, social justice, and joyful commitment. *Western Journal of Communication*, 74(1), 68-93.
- Hausman, A., Kabadayi, S., & Price, K. (2014). Consumer–brand engagement on Facebook: liking and commenting behaviors. *Journal of research in interactive marketing*.
- Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*. Guilford Publications.
- Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal (AMJ)*, 25(1), 76-81.
- Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling*, 17(1), 82-109.
- Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities. *Journal of Marketing*, 0022242918822300.
- Hess Jr, R. L., Ganesan, S., & Klein, N. M. (2003). Service failure and recovery: the impact of relationship factors on customer satisfaction. *Journal of the Academy of Marketing Science*, 31(2), 127-145.
- Hewett, K., Rand, W., Rust, R. T., & van Heerde, H. J. (2016). Brand buzz in the echoverse. *Journal of Marketing*, 80(3), 1-24.
- Hodges, B., Warren, C., & Estes, Z. (2016). More Than Words: a Psycholinguistic Perspective on the Properties of Effective Brand Slogans. *ACR North American Advances*.

- Holman, M. R., & Lay, J. C. (2019). They see dead people (voting): Correcting misperceptions about voter fraud in the 2016 US presidential election. *Journal of Political Marketing*, 18(1-2), 31-68.
- Hollenbeck, C. R., & Zinkhan, G. M. (2006). Consumer activism on the internet: The role of anti-brand communities. *ACR North American Advances*.
- Holt, D. B. (2002). Why do brands cause trouble? A dialectical theory of consumer culture and branding. *Journal of Consumer Research*, 29(1), 70-90.
- Horvath, B. (2014). Reputations, Trolls, and the Internet. Retrieved from (<https://www.business2community.com/crisis-management/reputations-trolls-internet-0860269>)
- Howard, J. A., & Sheth, J. N. (1969). The Theory of Buyer Behavior (No. 658.834 H6).
- Hsiao, C. H., Shen, G. C., & Chao, P. J. (2015). How does brand misconduct affect the brand–customer relationship?. *Journal of Business Research*, 68(4), 862-866
- Hulland, J., Baumgartner, H., & Smith, K. M. (2018). Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. *Journal of the Academy of Marketing Science*, 46(1), 92-108.
- Hulland, J., & Miller, J. (2018). Keep on Turkin'? *Journal of the Academy of Marketing Science* (46), 789–794
- Hunt, H. K. (1973). Effects of corrective advertising. *Journal of Advertising Research*.
- Ingold, P. V., Dönni, M., & Lievens, F. (2018). A dual-process theory perspective to better understand judgments in assessment centers: The role of initial impressions for dimension ratings and validity. *The Journal of Applied Psychology*.
- Inman, J. J., & Zeelenberg, M. (2002). Regret in repeat purchase versus switching decisions: The attenuating role of decision justifiability. *Journal of Consumer Research*, 29(1), 116-128.
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218.
- Jayasimha, K. R., Chaudhary, H., & Chauhan, A. (2017). Investigating consumer advocacy, community usefulness, and brand avoidance. *Marketing Intelligence & Planning*, 35(4), 488-509.
- Johnson, A. R., Matear, M., & Thomson, M. (2010). A coal in the heart: Self-relevance as a post-exit predictor of consumer anti-brand actions. *Journal of Consumer Research*, 38(1), 108-125.

- Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical research memoirs*.
- Johnson, B. & Smith, D.G. (2017). Too Many Men Are Silent Bystanders to Sexual Harassment. Harvard Business Review Online. Retrieved from (<https://hbr.org/2017/03/too-many-men-are-silent-bystanders-to-sexual-harassment>)
- Jonason, P. K., Duineveld, J. J., & Middleton, J. P. (2015). Pathology, pseudopathology, and the Dark Triad of personality. *Personality and Individual Differences*, 78, 43-47.
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the Dark Triad. *Psychological Assessment*, 22(2), 420.
- Jonason, P. K., Li, N. P., Webster, G. D., & Schmitt, D. P. (2009). The Dark Triad: Facilitating a short-term mating strategy in men. *European Journal of Personality: Published for the European Association of Personality Psychology*, 23(1), 5-18.
- Jonason, P. K., & Tost, J. (2010). I just cannot control myself: The Dark Triad and self-control. *Personality and Individual Differences*, 49(6), 611-615.
- Jones, D. N., & Paulhus, D. L. (2014). Introducing the short Dark Triad (SD3) a brief measure of dark personality traits. *Assessment*, 21(1), 28-41.
- Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33(2), 125-143.
- Kahneman, D., & Egan, P. (2011). *Thinking, Fast and Slow* (Vol. 1). New York: Farrar, Straus and Giroux.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The Psychology of Intuitive Judgment*, 49, 81.
- Kalyanam, K., McIntyre, S., & Masonis, J. T. (2007). Adaptive experimentation in interactive marketing: The case of viral marketing at Plaxo. *Journal of Interactive Marketing*, 21(3), 72-85.
- Kane, J. E., & Webster, G. D. (2012). Heuristics and Biases That Help and Hinder Scientists: Toward a Psychology of Scientific Judgment and Decision Making. *Handbook of the Psychology of Science*, 437.
- Karampournioti, E., Hennigs, N., & Wiedmann, K. P. (2018). When pain is pleasure: Identifying consumer psychopaths. *Psychology & Marketing*, 35(4), 268-282.

- Kennedy, E. J., Lawton, L., & Plumlee, E. L. (2002). Blissful ignorance: The problem of unrecognized incompetence and academic performance. *Journal of Marketing Education*, 24(3), 243-252.
- Khamitov, M., Grégoire, Y., & Suri, A. (2019). A systematic review of brand transgression, service failure recovery and product-harm crisis: integration and guiding insights. *Journal of the Academy of Marketing Science*, 1-24.
- Kim, S. (2013). Does corporate advertising work in a crisis? An examination of inoculation theory. *Journal of Marketing Communications*, 19(4), 293-305.
- Kim, K. J., Kim, Y. H., & Kim, M. S. A Study on Design Information Search Process of Consumer Behavior in Online Community.
- Kim, J., Kim, J. E., & Park, J. (2012). Effects of cognitive resource availability on consumer decisions involving counterfeit products: The role of perceived justification. *Marketing Letters*, 23(3), 869-881.
- Kim, S., & Rucker, D. D. (2012). Bracing for the psychological storm: Proactive versus reactive compensatory consumption. *Journal of Consumer Research*, 39(4), 815-830.
- Koh, S. C. L., Gunasekaran, A., & Rajkumar, D. (2008). ERP II: The involvement, benefits and impediments of collaborative information sharing. *International Journal of Production Economics*, 113(1), 245-268.
- Kolyesnikova, N., Sullivan Dodd, S. L., & Callison, C. (2011). Consumer affective responses to direct mail messages: The effect of gratitude and obligation. *Journal of Marketing Communications*, 17(5), 337-353.
- Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, 39(1), 61-72.
- Kozinets, R. V. (2015). *Netnography: redefined*. Sage.
- Kozinets, R. V. (2016). Amazonian forests and trees: Multiplicity and objectivity in studies of online consumer-generated ratings and reviews, a commentary on de Langhe, Fernbach, and Lichtenstein. *Journal of Consumer Research*, 42(6), 834-839.
- Kravets, O. (2007). Consumer Vigilantism. *ACR North American Advances*.
- Krishnamurthy, S., & Kucuk, S. U. (2009). Anti-branding on the internet. *Journal of Business Research*, 62(11), 1119-1126.
- Kucuk, S. U. (2008). Negative double jeopardy: The role of anti-brand sites on the internet. *Journal of Brand Management*, 15(3), 209-222.

- Kucuk, S. U. (2010). Negative double jeopardy revisited: A longitudinal analysis. *Journal of Brand Management*, 18(2), 150-158.
- Kuran, T., & Sunstein, C. R. (1998). Availability cascades and risk regulation. *Stan. L. Rev.*, 51, 683.
- Kwok, L., & Yu, B. (2013). Spreading social media messages on Facebook: An analysis of restaurant business-to-consumer communications. *Cornell Hospitality Quarterly*, 54(1), 84-94.
- Langer, R., & Beckman, S. C. (2005). Sensitive research topics: netnography revisited. *Qualitative Market Research: An International Journal*, 8(2), 189-203.
- Langner, S., Hennigs, N., & Wiedmann, K. P. (2013). Social persuasion: targeting social identities through social influencers. *Journal of Consumer Marketing*, 30(1), 31-49.
- Larson, L. R., & Denton, L. T. (2014). eWOM Watchdogs: Ego-Threatening Product Domains and the Policing of Positive Online Reviews. *Psychology & Marketing*, 31(9), 801-811.
- Lee, N., & Cadogan, J. W. (2013). Problems with formative and higher-order reflective variables. *Journal of Business Research*, 66(2), 242-247.
- Lee, C. S., & Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. *Computers in human behavior*, 28(2), 331-339.
- Lee, M. S., & Male, M. (2011). Against medical advice: the anti-consumption of vaccines. *Journal of Consumer Marketing*.
- Leibel, W. (1991). When scientists are wrong: admitting inadvertent error in research. *Journal of Business Ethics*, 601-604.
- Lessne, G. J., & Didow Jr, N. M. (1987). Inoculation theory and resistance to persuasion in marketing. *Psychology & Marketing*, 4(2), 157-165.
- Lewis, I., Watson, B., White, K. M., & Elliott, B. (2013). The beliefs which influence young males to speed and strategies to slow them down: informing the content of antispeeding messages. *Psychology & Marketing*, 30(9), 826-841.
- Levine, T. R. (2014). Truth-default theory (TDT) a theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4), 378-392.
- Levine, M., & Crowther, S. (2008). The responsive bystander: How social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of personality and social psychology*, 95(6), 1429.
- Levitin, D. J. (2016). *A field guide to lies: Critical thinking in the information age*. Penguin.

- Levy, P., Lundgren, D., Ansel, M., Fell, D., Fink, B., & McGrath, J. E. (1973). "Bystander Effect in a Demand-without-Threat Situation": Erratum.
- Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, 25(4), 361-381.
- Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2012). Structural equation modeling of mediation and moderation with contextual factors. In *Modeling Contextual Effects in Longitudinal Studies* (pp. 211-234). Routledge.
- Lizardo, O., Mowry, R., Sepulvado, B., Stoltz, D. S., Taylor, M. A., Van Ness, J., & Wood, M. (2016). What are dual process models? Implications for cultural analysis in sociology. *Sociological Theory*, 34(4), 287-310.
- Loeb, H. (2019) February. *Business Must Combat Fake News*. Retrieved from <https://www.edelman.com/insights/business-must-combat-fake-news>
- Lomax, R. G., & Schumacker, R. E. (2004). *A beginner's guide to structural equation modeling*. psychology press.
- Ludwig, S., De Ruyter, K., Friedman, M., Brügggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1), 87-103.
- Machackova, H., Dedkova, L., & Mezulanikova, K. (2015). Brief report: The bystander effect in cyberbullying incidents. *Journal of adolescence*, 43, 96-99.
- Malaviya, P., & Sivakumar, K. (2002). The influence of choice justification and stimulus meaningfulness on the attraction effect. *Journal of Marketing Theory and Practice*, 10(4), 20-29.
- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320-341.
- Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*, 39(1), 87-96.
- McCornack, S.A., Levine, T.R., Solowczuk, K.A., Torres, H.I. & Campbell, D.M. (1992). When the alteration of information is viewed as deception: An empirical test of information manipulation theory. *Communication Monographs*, 59(1), 17-30.

- McGuire, W. J. (1964). Some contemporary approaches. In *Advances in Experimental Social Psychology* (Vol. 1, pp. 191-229). Academic Press.
- McGuire, W. J. (1968). Personality and attitude change: An information processing theory. In A. G. Greenwald, T. C. Brock, and T. M. Ostrom (eds.), *Psychological Foundations of Attitudes* (pp. 171-196). San Diego, CA: Academic Press.
- Meddaugh, P. (2010). Bakhtin, Colbert, and the Center of Discourse: Is There No “Truthiness in Humor? *Critical Studies in Media Communication*, 27 (4), 376-390.
- Mehl, M. R. (2006). Quantitative text analysis. *Handbook of Multimethod Measurement in Psychology*, 141-156.
- Meschke, M. (2014) June. *Internet Trolls: Do they hurt your business?* Retrieved from <https://www.michigancreative.com/blog/internet-trolls-do-they-hurt-your-business/>
- Meyers-Levy, J., & Maheswaran, D. (2004). Exploring message framing outcomes when systematic, heuristic, or both types of processing occur. *Journal of Consumer psychology*, 14(1-2), 159-167.
- Mikolon, S., Quaiser, B., & Wieseke, J. (2015). Don't try harder: using customer inoculation to build resistance against service failures. *Journal of the Academy of Marketing Science*, 43(4), 512-527.
- Mittal, V. (2015). Qualitative Research for Customer-Focused Insights. Teaching Note, Rice University, Houston, TX 77005.
- Moore, D., Shannon, C., Voelker, G. M., & Savage, S. (2003, March). Internet quarantine: Requirements for containing self-propagating code. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)* (Vol. 3, pp. 1901-1910). IEEE.
- Moschis, G. P., & Cox, D. (1989). Deviant consumer behavior. *ACR North American Advances*.
- Munger, M. (2008). Blogging and political information: truth or truthiness? *Public Choice*, 134, 125-138
- Muniz, A. M., & O'guinn, T. C. (2001). Brand community. *Journal of Consumer Research*, 27(4), 412-432.
- Muñiz, Jr, A. M., & Schau, H. J. (2007). Vigilante marketing and consumer-created communications. *Journal of Advertising*, 36(3), 35-50.
- Munnukka, J., Uusitalo, O., & Toivonen, H. (2016). Credibility of a peer endorser and advertising effectiveness. *Journal of Consumer Marketing*, 33(3), 182-192.

- Murray, S. L., Pinkus, R. T., Holmes, J. G., Harris, B., Gomillion, S., Aloni, M., ... & Leder, S. (2011). Signaling when (and when not) to be cautious and self-protective: Impulsive and reflective trust in close relationships. *Journal of Personality and Social Psychology*, *101*(3), 485.
- Nazlan, N. H., Tanford, S., & Montgomery, R. (2018). The effect of availability heuristics in online consumer reviews. *Journal of Consumer Behaviour*, *17*(5), 449-460.
- Nelson, M. R., & Otnes, C. C. (2005). Exploring cross-cultural ambivalence: a netnography of intercultural wedding message boards. *Journal of Business Research*, *58*(1), 89-95.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Sage Publications.
- Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, *19*(5), 969-974.
- Newman, M.L., Pennebaker, J.W., Berry, D.S., & Richards, J.M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*, 665-675.
- Nunnally, J.C., Bernstein, I.H., 1994. *Psychometric Theory*, 3rd edition. McGraw-Hill, Inc, New York
- Ognyanova, K. (2019). The social context of media trust: A network influence model. *Journal of Communication*, *69*(5), 544-567.
- Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of marketing research*, *42*(1), 43-53.
- O'meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological assessment*, *23*(2), 523.
- O'Sullivan, S. R. (2015). The market maven crowd: Collaborative risk-aversion and enhanced consumption context control in an illicit market. *Psychology & Marketing*, *32*(3), 285-302.
- Pagani, M., Hofacker, C. F., & Goldsmith, R. E. (2011). The influence of personality on active and passive use of social networking sites. *Psychology & Marketing*, *28*(5), 441-456.
- Pantelidis, I. S. (2010). Electronic meal experience: A content analysis of online restaurant comments. *Cornell Hospitality Quarterly*, *51*(4), 483-491.

- Papageorgis, D., and McGuire, W. J. (1961). The generality of immunity to persuasion produced by pre-exposure to weakened counterarguments. *Journal of Abnormal and Social Psychology*, 62, 475-481.
- Park, H., & Cho, H. (2012). Social network online communities: information sources for apparel shopping. *Journal of Consumer Marketing*, 29(6), 400-411.
- Parrott, R., Volkman, J., Ghetian, C., Weiner, J., Raup-Krieger, J., & Parrott, J. (2008). Memorable messages about genes and health: Implications for direct-to-consumer marketing of genetic tests and therapies. *Health marketing quarterly*, 25(1-2), 8-32.
- Parry, W. (2011) March. *How to Spot Psychopaths: Speech Patterns Give Them Away*. Retrieved from <https://www.livescience.com/16585-psychopaths-speech-language.html>
- Payne, A., & Frow, P. (2017). Relationship marketing: looking backwards towards the future. *Journal of services marketing*, 31(1), 11-15.
- Peisley, T. (2017) March. *Is narcissism common? The answer may surprise you*. Retrieved from <https://www.sane.org/the-sane-blog/mental-illness/is-narcissism-common-the-answer-may-surprise-you>
- Pennebaker, J.W. (2019) February. *Interpreting LIWC Output*. Retrieved from <https://liwc.wpengine.com/interpreting-liwc-output/>
- Pennebaker, J.W. (2011). *The Secret Life of Pronouns: What Our Words Say About Us* (NY: Bloomsbury).
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PloS one*, 9(12), e115844.
- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2), 381-391.
- Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, 67(5), 1035-1041.
- Pfau, M. (1997). The inoculation model of resistance to influence. In F.J. Boster & G. Barnett (Eds.), *Progress in Communication Sciences* (Vol. 13, pp. 133–171). Norwood, NJ: Ablex.
- Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications*, 20(1-2), 117-128.

- Phan, T. Q., & Godes, D. (2018). The evolution of influence through endogenous link formation. *Marketing Science*, 37(2), 259-278.
- Pickard-Whitehead, G. (2017) July. *What is an Internet Troll and How Can One Ruin Your Business?* Retrieved from <https://smallbiztrends.com/2017/07/what-is-an-internet-troll.html>
- Pitt, L. F., Ewing, M. T., & Berthon, P. R. (2002). Proactive behavior and industrial salesforce performance. *Industrial marketing management*, 31(8), 639-644.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879
- Pollock, T. G., Rindova, V. P., & Maggitti, P. G. (2008). Market watch: Information and availability cascades among the media and investors in the US IPO market. *Academy of Management Journal*, 51(2), 335-358.
- Pongsakornrunsilp, S., & Schroeder, J. E. (2011). Understanding value co-creation in a co-consuming brand community. *Marketing Theory*, 11(3), 303-324.
- Porter, S., Bhanwer, A., Woodworth, M., & Black, P. J. (2014). Soldiers of misfortune: An examination of the Dark Triad and the experience of schadenfreude. *Personality and Individual Differences*, 67, 64-68.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717-731.
- Rains, S. A. (2007). The impact of anonymity on perceptions of source credibility and influence in computer-mediated group communication: A test of two competing hypotheses. *Communication research*, 34(1), 100-125.
- Rampl, L. V., Opitz, C., Welp, I. M., & Kenning, P. (2016). The role of emotions in decision-making on employer brands: insights from functional magnetic resonance imaging (fMRI). *Marketing Letters*, 27(2), 361-374.
- Rapp-Ricciardi, M., Widh, J., Barbieri, B., Amato, C., & Archer, T. (2018). Dark Triad, Locus of Control and Affective Status Among Individuals with an Entrepreneurial Intent. *Journal of Entrepreneurship Education*, 21(1).
- Rauschnabel, P. A., Kammerlander, N., & Ivens, B. S. (2016). Collaborative brand attacks in social media: Exploring the antecedents, characteristics, and consequences of a new form of brand crises. *Journal of Marketing Theory and Practice*, 24(4), 381-410.

- Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, 81(12), 46-55.
- Render, J. (2018). *Why consumers trust each other more than brands*. Retrieved from <https://blog.consumeraffairs.com/why-consumers-trust-each-other-more-than-brands-908a440ed8d7>
- Richardson, L., & St Pierre, E. (2008). A method of inquiry. *Collecting and Interpreting Qualitative Materials*, 3(4), 473.
- Roberts, D. (2016) November. *Post-truth politics*. Accessed from <http://grist.org/article/2010-0330-post-truth-politics/>
- Rosenbaum, M. S., Kuntze, R., & Wooldridge, B. R. (2011). Understanding unethical retail disposition practice and restraint from the consumer perspective. *Psychology & Marketing*, 28(1), 29-52.
- Ro, H., & Wong, J. (2012). Customer opportunistic complaints management: A critical incident approach. *International Journal of Hospitality Management*, 31(2), 419-427.
- Salmivalli, C. (2014). Participant roles in bullying: How can peer bystanders be utilized in interventions?. *Theory Into Practice*, 53(4), 286-292.
- Sato, A. (2009). A study of export expansion of japanese fruits toward Taiwan market: In connection with export-import quarantine system. *Agricultural Marketing Journal of Japan (Japan)*.
- Saucier, D. A., & Webster, R. J. (2010). Social vigilantism: Measuring individual differences in belief superiority and resistance to persuasion. *Personality and Social Psychology Bulletin*, 36(1), 19-32.
- Saucier, D. A., Webster, R. J., Hoffman, B. H., & Strain, M. L. (2014). Social vigilantism and reported use of strategies to resist persuasion. *Personality and Individual Differences*, 70, 120-125.
- Schau, H. J., Muñiz Jr, A. M., & Arnould, E. J. (2009). How brand community practices create value. *Journal of marketing*, 73(5), 30-51.
- Scholz, J., & Smith, A. N. (2019). Branding in the age of social media firestorms: how to create brand value by fighting back online. *Journal of Marketing Management*, 1-35.
- Schweitzer, M., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty*, 25(2), 185–201.

- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. psychology press. 2nd edition
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology, 14*(4), 332-348.
- Seibert, S. E., Kraimer, M. L., & Crant, J. M. (2001). What do proactive people do? A longitudinal model linking proactive personality and career success. *Personnel Psychology, 54*(4), 845-874.
- Schau, H. J., Muñiz Jr, A. M., & Arnould, E. J. (2009). How brand community practices create value. *Journal of Marketing, 73*(5), 30-51.
- Schivinski, B., & Dabrowski, D. (2016). The effect of social media communication on consumer perceptions of brands. *Journal of Marketing Communications, 22*(2), 189-214.
- Sen, S., Gürhan-Canli, Z., & Morwitz, V. (2001). Withholding consumption: A social dilemma perspective on consumer boycotts. *Journal of Consumer research, 28*(3), 399-417.
- Sharpe Wessling, K., Huber, J., & Netzer, O. (2017). MTurk character misrepresentation: Assessment and solutions. *Journal of Consumer Research, 44*(1), 211-230.
- Shaw-Ching Liu, B., Sudharshan, D., & Hamer, L. O. (2000). After-service response in service quality assessment: a real-time updating model approach. *Journal of Services Marketing, 14*(2), 160-177.
- Sonnentag, T. L., & Barnett, M. A. (2016). Role of moral identity and moral courage characteristics in adolescents' tendencies to be a moral rebel. *Ethics & Behavior, 26*(4), 277-299.
- Sparks, J. R. (1994). Machiavellianism and personal success in marketing: The moderating role of latitude for improvisation. *Journal of the Academy of Marketing Science, 22*(4), 393.
- Spears, N. (2006). Just moseying around and happening upon it versus a master plan: Minimizing regret in impulse versus planned sales promotion purchases. *Psychology & Marketing, 23*(1), 57-73.
- Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research, 50*(2), 277-288.
- Stokburger-Sauer, N. E., & Wiertz, C. (2015). Online consumption communities: An introduction. *Psychology & Marketing, 32*(3), 235-239.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior, 7*(3), 321-326.

- Sun, L. G. (2012). Disaster mythology and availability cascades. *Duke Env'tl. L. & Pol'y F.*, 23, 73.
- Szybillo, G. J., & Heslin, R. (1973). Resistance to persuasion: Inoculation theory in a marketing context. *Journal of Marketing Research*, 396-403.
- Tausczik, Y.R., & Pennebaker, J.W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24-54.
- Thomas, Rob. (2013). How One Business Saw Its Reputation Destroyed Within 24 Hours. And How You Can Stop it Happening to You. Retrieved from (<http://blog.wsi-emarketing.com/how-one-business-saw-its-reputation-destroyed-within-24-hours-and-how-you-can-stop-it-happening-to-you/>)
- Thyroff, A., & Kilbourne, W. E. (2018). Self-enhancement and individual competitiveness as mediators in the materialism/consumer satisfaction relationship. *Journal of Business Research*, 92, 189-196.
- Tseng, S. Y., & Wang, C. N. (2016). Perceived risk influence on dual-route information adoption processes on travel websites. *Journal of Business Research*, 69(6), 2289-2296.
- Tsujii, T., & Watanabe, S. (2009). Neural correlates of dual-task effect on belief-bias syllogistic reasoning: a near-infrared spectroscopy study. *Brain Research*, 1287, 118-125.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2), 207-232.
- Underwood, J. D., Kerlin, L., & Farrington-Flint, L. (2011). The lies we tell and what they say about us: Using behavioural characteristics to explain Facebook activity. *Computers in Human Behavior*, 27(5), 1621-1626.
- Veil, S. R., & Kent, M. L. (2008). Issues management and inoculation: Tylenol's responsible dosing advertising. *Public Relations Review*, 34(4), 399-402.
- Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2), 231.
- Visentin, M., Pizzi, G., & Pichierri, M. (2019). Fake News, Real Problems for Brands: The Impact of Content Truthfulness and Source Credibility on consumers' Behavioral Intentions toward the Advertised Brands. *Journal of Interactive Marketing*, 45, 99-112.
- Wakefield, K. L., & Inman, J. J. (1993). Who are the price vigilantes? An investigation of differentiating characteristics influencing price information processing. *Journal of Retailing*, 69(2), 216-233.

- Wallace, E., Buil, I., & de Chernatony, L. (2014). Consumer engagement with self-expressive brands: brand love and WOM outcomes. *Journal of Product & Brand Management*, 23(1), 33-42.
- Werle, C. O., & Cuny, C. (2012). The boomerang effect of mandatory sanitary messages to prevent obesity. *Marketing Letters*, 23(3), 883-891.
- Wirtz, J., Den Ambtman, A., Bloemer, J., Horváth, C., Ramaseshan, B., Van De Klundert, J., ... & Kandampully, J. (2013). Managing brands and customer engagement in online brand communities. *Journal of Service Management*, 24(3), 223-244.
- Wirtz, J., & Kum, D. (2004). Consumer cheating on service guarantees. *Journal of the Academy of Marketing Science*, 32(2), 159-175.
- Wisse, B., Barelds, D. P., & Rietzschel, E. F. (2015). How innovative is your employee? The role of employee and supervisor Dark Triad personality traits in supervisor perceptions of employee innovative behavior. *Personality and Individual Differences*, 82, 158-162.
- Wolfgang, J. D. (2018). Taming the ‘trolls’: How journalists negotiate the boundaries of journalism and online comments. *Journalism*, 1464884918762362.
- Woodcock, A. (2019). Nearly half of social media users who share articles have passed on fake news, study suggests. Retrieved from (<https://www.independent.co.uk/news/uk/home-news/fake-news-facebook-twitter-share-misinformation-survey-a8908361.html>)
- Wu, Y., Wen, Z., Marsh, H. W., & Hau, K. T. (2013). A comparison of strategies for forming product indicators for unequal numbers of items in structural equation models of latent interactions. *Structural equation modeling: A multidisciplinary journal*, 20(4), 551-567.
- Yang, Z., & Peterson, R. T. (2003). I read about it online... *Marketing Research*, 15(4).
- Yoo, B., & Donthu, N. (2001). Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research*, 52(1), 1-14.
- Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197-206.
- Zhu, R., Dholakia, U. M., Chen, X., & Algesheimer, R. (2012). Does online community participation foster risky financial behavior?. *Journal of Marketing Research*, 49(3), 394-407.
- Zolfagharian, M., & Yazdanparast, A. (2017). The dark side of consumer life in the age of virtual and mobile technology. *Journal of Marketing Management*, 33(15-16), 1304-1335.

Zou, C. C., Gong, W., & Towsley, D. (2003, October). Worm propagation modeling and analysis under dynamic quarantine defense. In *Proceedings of the 2003 ACM workshop on Rapid malware* (pp. 51-60).

APPENDIX A

ESSAY ONE

Memos

1. Present – 12/2018

Theory: Inoculation Theory - Dual-process model Information Manipulation Theory, Theory of Planned Behavior

The need for power that was originally posited as a motivator of the negative online behavior was reexamined by viewing other forums and posts. It appeared that power is a somewhat small explanation for the phenomenon. For instance, many people that post negative information may post for other reasons such as opportunism. This explanation can encompass the need for power in that the poster finds an opportunity to exploit others to address that need. The same pursuit of possibilities can be linked to financial gains, humor, and manipulation of others. This pursuit also appears to be accompanied by varying levels of vigilantism based on the online environment. For example, the online disinhibition effect (discussed in the forums) states that people use an online environment differently from their real life. Respondents can use online channels to facilitate a different persona since the platforms provide levels of anonymity and a disconnect. This effect seems to strengthen the likelihood of the person using online forums as a platform to coordinate and incite others. Interestingly, a participant can utilize the leverage of the group to meet one's ends. When the participant is likely to use the forum to act as a vigilante for sharing information and creating a platform, he/she can then engage in malicious behavior through the creation of fraudulent or misleading information. This likelihood can be intensified when the poster finds justification for the sadistic action (i.e., punishing or harming another). These constructs tend to explain the discussions taking place in the online forum that focused on the pricing failure. The concepts should be further explored since consumers can easily build forums to address a specific need then disband afterward. Additionally, potential customers can find both current and dissolved discussions. This can act to potentially harm a company well after a crisis has been addressed.

2. Initial Exposure – 02/2016

Memo: I was informed of a pricing error that was being discussed by participants on Reddit. An individual that was aware of the situation directed me to the page. At first glance, I found the entire situation amusing (similar to the participants in the forum). I looked specifically for humorous comments that typically referred to numerous companies in a negative light. Comments voiced contempt toward the companies that would not honor pricing. After I browsed through multiple "humorous" comments, I began to see additional pieces of information that were interesting. I had not started a doctoral seminar, nor was I familiar with any research processes to analyze the data, but the topics stuck with me. The respondents began to coordinate with one another based on the information provided by the numerous companies involved. The forum began to work as a unit that updated in real-time. Purchasers shared whether or not the pricing was honored, timelines are given, and options for fulfillment. The forum suggested the communities can be formed quickly for a given purpose and address a core need quite swiftly. Additionally, the respondents began to coordinate negative reviews, communicated detailed information for online marketplace policies, and penalties that the companies could experience. The customers were well informed, tactical, and resilient in the pursuit of the forum's purpose, getting the pricing.

3. Updates from Promotion and Distribution Seminar – 10/2016

Memo: Fake news and Truthiness

Theory: Social Threshold Theory

The 2016 election influenced my second assessment of the online forum. Much of the political climate revolved around the phrase "fake news." I was thinking more about how fraudulent information can be disseminated online so easily. I thought about the term "truthiness" coined by Stephen Colbert that addresses information that just "seems" correct, reinforces current beliefs, and is typically accepted as fact by individuals. Social media serves a crucial purpose for communication between customers and brands. However, what can occur when one party purposely attends to persuade and dissuade others? I returned to the original forum to compare the purposeful sharing of fraudulent information. I think that the online forum also relied on this type of approach to some extent. For instance, customers voiced discontent and shared information that was purported to come from the companies. However, little verification took place. The customers tended to believe the information that was posted (i.e., when a customer claimed that the order was delivered or the pricing was honored). This hinted that there was some confirmation bias at play in the community. The customers were focused on getting the pricing at all costs. I believe that some influence or bias influences whether or not a customer believes others or the company.

See excerpt from the paper:

In today's market, very little is done to regulate messages shared on social media. Communication campaigns today rely on social media to help messages spread. Unfortunately, many companies can lose control of such programs. This paper aims to develop the concept of truthiness further and determine the particular social and psychological elements that attribute to the continuation of this phenomenon in advertising and on social media. The social threshold is examined to understand the relationship between motivated reasoning and acceptance of truth in the messages. (Granovetter, 1978) (Hasan and Ukkusuri, 2011). Readers may seek to confirm these motivated biases rather than factual information (Roberts, 2016) (Moore, Harris, and Chen, 1995). This motivated reasoning mentality, when moderated by the social threshold level, can then lead to message acceptance and propensity to share the message with others. By sharing the message, the cycle continues as more members of the sender's social network are exposed to misleading information. The purpose of this paper is to create the initial framework for this phenomenon by defining the constructs of interest, discussing potential relationships, and creating the conceptual framework for the model.

4. Updates from Qualitative Seminar – 08/2017

Memo: Anti-Brand Communities

Theory: Grounded Theory

In the qualitative methods seminar, I was very interested in doing more analysis into the types of sentiments of those in the forum. I was also interested in the idea of anti-brand communities that focused on destroying brands. For example, as customers can build communities through camaraderie and connection with a brand, customers can build communities against companies or brands that do not reinforce an identity. Moreover, customers can act to fact check companies and advocate for other customers. Vigilantism is first addressed here as a customer seeking out

information online to help other consumers. When a company acts unjustly or wrongly, these customers act to help others by bringing fraudulent information to light. It seems that the participants had motivations beyond just the incentive for a low price. Some participants not only wanted to take advantage of the price but also coordinate discussions and educate others. These customers focused on updates based on updates in company response. It can also be used for more exploitative purposes through coordinated efforts to exploit pricing mistakes. So, with enough support and leverage, the group can have sway over the companies involved in the pricing error.

See excerpt from the paper:

Consumer groups or communities are common in both online and offline settings. These groups can be used to educate fellow consumers, discuss product capabilities, and facilitate product usage. Also, these groups can act as a community for product enthusiasts. The message boards, groups, comments, and reviews provided within these communities deliver invaluable insights to firms that hope to understand the sentiments and perceptions of their consumers. Online communities can be facilitated by the firm to provide a positive and productive environment for the sharing of experiences. However, consumer groups can form independently to coordinate reviews, word of mouth, and pricing. These groups and forums can become vigilant in monitoring fairness in pricing and offerings. When a brand does not manage expectations effectively or fairly, these groups can become a breeding ground for dissonance and vigilance. The information gained by users of the groups can then be used to dictate future behavior and lead to negative online reviews and coordinated brand destruction. In addition, these groups can be highly collaborative and look to exploit pricing or mistakes made by a seller. Firms should be diligent in managing the online image of their brand and be mindful of the sentiments of consumers. Firms should monitor the types of conversations that take place in online consumer groups, determine the level of vigilance associated with the group (through the comparison of pricing, complaint resolution, return policies, guarantees, and promises), and determine the role of perceived fairness and distributive justice in mitigating perceived failures in product or service offerings. By addressing these concerns, it is hoped that firms may address possible distress sooner to alleviate the threat of negative word of mouth, brand destruction, loss of sales, and negative online reviews.

5. Updates from Consumer Behavior Seminar – 10/2017

Memo: Betrayal

Theory: Attribution Theory and Justice Theory

I began to question the intentions of those involved in the forum and wondered if maybe the retaliation was more of a response to a violation of trust. For instance, betrayal focuses on a relational party violating relational norms and causing harm. Could the norms of the customers and the companies be broken by not honoring pricing? Are there explicit and implicit norms guiding relationships? I wondered if, by comparing the customer service responses, the customers were tipped off to perceived violations of trust. Many customers were upset that they had been lied to and felt like the companies were not forthcoming with information. Some stated that if it was a pricing mistake, admit the error and do not lie to the customers. Interestingly, the customers did not appear to have any previous interactions with the companies, and therefore

prior relationships did not seem to exist. It still seemed that the customers felt betrayed by the actions of the companies and thus sought to retaliate afterward.

See excerpt from the paper:

The concept of betrayal has been studied in both B2B and B2C contexts as a means to address the influence and severity of relational norms violations in exchange transactions (Finkel et al., 2007). Prior research has suggested the importance of addressing betrayal as an outcome of a service failure. By mitigating the effects of betrayal through recovery, firms can limit the consumer's perception of the severity of the failure, likelihood to engage in negative word-of-mouth, and a desire for revenge (Gregoire and Fisher, 2008). Although prior betrayal research has suggested the importance of relational norm components in setting consumer expectations, little research has attempted to develop a categorization of the types of betrayal that arise between consumers and businesses.

6. Second-year paper review – 02/2018

Memo: Jealousy and Envy

Theory: Equity Theory

I began to think about the core drivers of online communications and the role of comparisons to others. Betrayal can occur when a relational party feels taken advantage of. However, what drives the betrayal? I found literature that discussed the concepts of jealousy and envy in consumer settings. For instance, jealousy arises when a customer feels that a relational partner has transferred attention to another customer. This dynamic suggests that three parties are involved. Envy occurs when a customer desires the possessions or attributes of another customer. The conversations being had tended to focus on the actions of the company rather than the possessions of the other customers. This was interesting since these customers had seemingly never used the companies before. Therefore, jealousy and envy can also influence the retaliation that took place. The customers shared a negative word of mouth and communicated about the negative reviews that were being posted. The customers also voiced contempt toward the customer service employees communicating updates. This action seemed to indicate that the customers were engaging in vindictive complaining to the company as well.

See excerpt from the paper:

Customers of all types are gaining greater access to other customers. This can be beneficial in developing communities associated with an offering, but companies can lose control over the flow of information. Because of this, customers can share information and then challenge the offerings given to others. This sharing of information can also be used as a means to monitor equity in the overall exchange relationship. For instance, salespeople, service providers, and even retailers aim to build different relationship types with customers. This especially important to address since both transactional and communal customers have expectations of norms that guide the relationships. When these customer treatments are compared and are deemed to deviate from norms, then the interaction may lead to adverse outcomes. By understanding the comparisons that are made and the role of equity in retaliation after a failure, firms can better equip themselves to handle customer comparisons in a variety of platforms.

7. Research Design Seminar – 5/2018

Memo: Hatred and Online Trolling

Theory: Dual Process Theory and Theory of Reasoned Action

What drives malicious online behavior, and are there any propensities that people may have that influence these tendencies? Does hatred toward a brand lead to an increased likelihood of acting maliciously? I began wondering if hatred can develop for brands online and with this hate came campaigns to discredit or harm companies. The Dark Triad personality traits (narcissism, Machiavellianism, and psychopathy) influence online "trolling" behavior. These components may indicate that the behaviors may have trait-based influences as well as motivations. If customers can act as advocates for brands and inform other customers, could these propensities be used in more negative ways? I think that customers can capitalize on the immediacy effect of information online. As someone posts more information and paints themselves as an expert, other customers can accept the message since access to external resources of information may be difficult. Therefore, when a customer hates a brand, it may vocalize the hate both explicitly through revenge-seeking or through more deceptive practices such as sabotage. Interestingly, a negative poster (by themselves) cannot do much damage to a company. These acts fall into some deviance that is pursued against a brand or company. However, when others are incentives, and the message takes on a critical mass, other customers may accept the information as fact.

See excerpt from the paper:

Consumer bystanders are exposed to numerous messages on a daily basis that can act to enhance or diminish attitudes toward a given brand. These messages can come from a place of benevolence through consumer and brand advocacy that provide information, benefits, and insights to other consumers. However, this information can come from a place of malice when consumers engage in insidious behaviors through malicious consumer deviance. Malicious consumer deviance is the act in which consumers act to mislead other consumers in forming attitudes toward a brand purposely. In these situations, the firm response is crucial in mitigating the effect of the falsified information. Both the firm and community response can act as a barrier to brand uncertainty and facilitate perceptions of trustworthiness and morality.

8. Casting a broader net to numerous Forums – 8/2018

Memo: Numerous Forums and Message Boards

Theory: Dual-process models, Information Manipulation Theory, Theory of Planned Behavior

I found it essential to ensure that the sentiments that were being voiced in the forum were not taking place in a bubble. I sought out numerous harmful or anti-brand websites, forums discussing trolling behavior, hated brands, unethical companies, and fake reviews. I read through 25 different sites or discussions to better determine the motivations and traits that influence the likelihood to engage in negative online behavior. I kept notes on the specific topics and the novel insights that emerged. I began to find themes associated with a need for power in driving online negative behavior (i.e., through trolling and negative information sharing). I also found discussions between those engaging in fake reviews and the readers of the forum. This venture provided greater insight into the motivations that arise. In addition, the forums had thousands of comments that showed much thought in the discussions. I then determined that the insights should be analyzed once again and then applied to the first forum to determine if the constructs could be implemented.

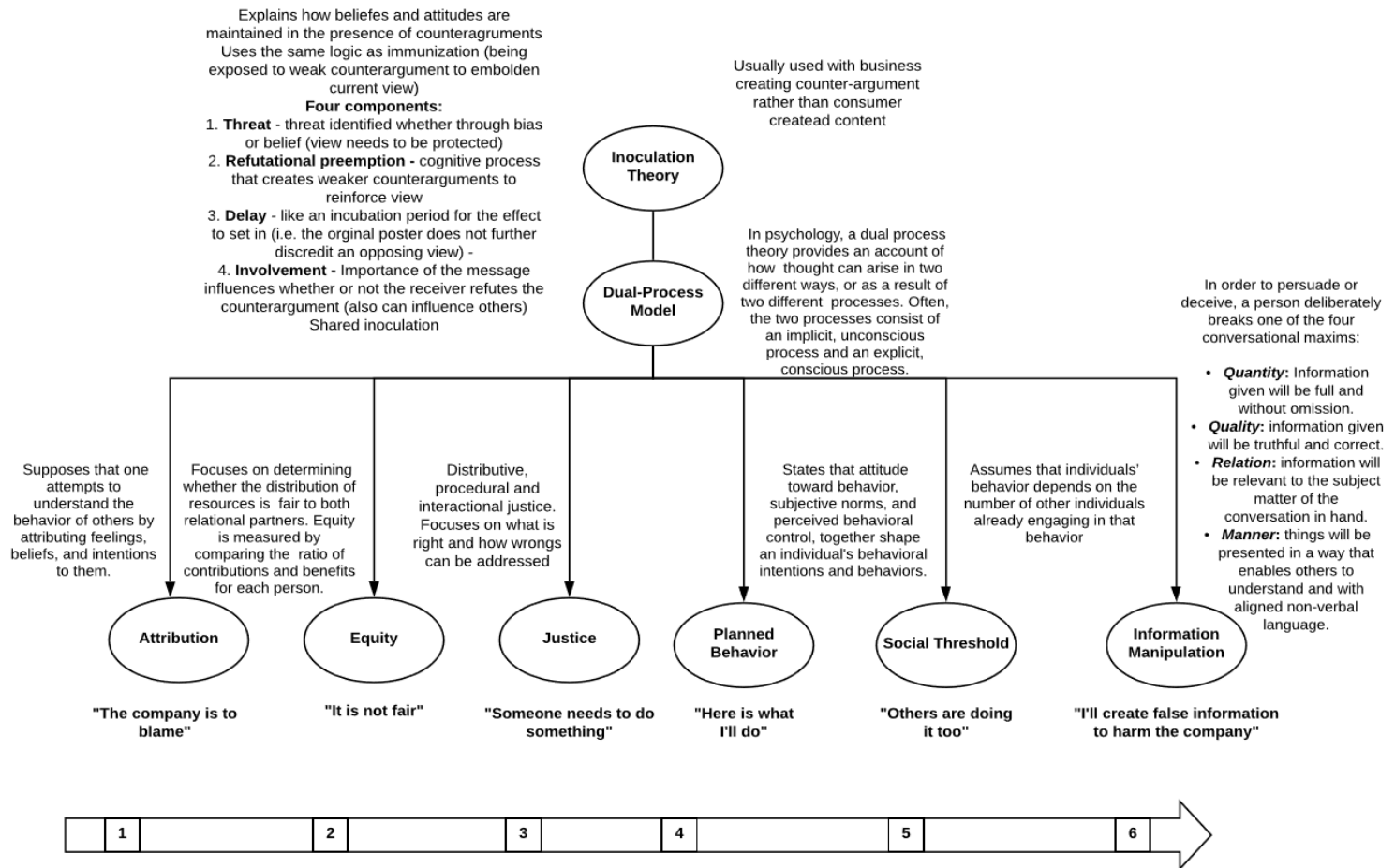


Figure A.1 Theoretical Development and Understanding

LIWC Output Overview

Retrieved from liwc.wpengine.com (Pennebaker, 2019)

I-words (I, Me, My) - percent of words in the post containing I, Me, and My

Social Words - percent of words in the post containing social words or group words

Positive Emotions - percent of words in the post containing positive emotion words

Negative Emotions - percent of words in the post containing negative emotion words

Cognitive Processes - percent of the words in the post that contain words

pertaining to cognitive thought (i.e. categories include insights, causation, discrepancy, tentative)

Summary Variables

LIWC also provides the summary variables analytical thinking, clout, authenticity, and emotional tone. The LIWC website and additional readings state that the variables are based on previous language research. The variables are based on standardized scores ranging from 0 to 100 (converted to percentiles).

Analytic

The analytical thinking variable is a factor-analytically derived dimension based on eight function word dimensions (c.f. Pennebaker, Chung, Frazee, Lavergne, & Beaver 2014). The analytic dimension addresses the degree to which people use words that suggest formal, logical, and hierarchical thinking patterns. People low in analytical thinking tend to write and think using language that is more narrative ways, focusing on the here-and-now, and personal experiences.

Clout

Clout refers to the relative social status, confidence, or leadership that people display through their writing or talking. The algorithm was developed based on the results from a series

of studies where people were interacting with one another (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2013)

Authenticity

Newman, Pennebaker, Berry, & Richards (2003) developed the algorithm for authenticity through a series of studies where people were induced to be honest or deceptive. In addition. Additionally, Pennebaker (2011) contributed to the creation of the algorithm through published summaries of deception studies. When people reveal themselves in an authentic or honest way, they are more personal, humble, and vulnerable (Pennebaker, 2011).

Emotional tone

The emotional tone variable puts both the positive and negative emotion dimensions into a summary variable (Cohn, Mehl, & Pennebaker, 2004). The algorithm is built so that the higher the number, the more positive the tone. The authors suggest that numbers below 50 suggest a more negative emotional tone (Cohn, Mehl, & Pennebaker, 2004).

Scales

Proactive Personality (Exploit Opportunity) (Bateman and Crant 1993)

1. I enjoy facing and overcoming obstacles to my ideas.
2. Nothing is more exciting than seeing my ideas turn into reality.
3. I excel at identifying opportunities.
4. I love to challenge the status quo.
5. I can spot a good opportunity long before others can.
6. I am great at turning problems into opportunities.
7. I am constantly looking for new ways to improve my life.
8. I feel driven to make a difference in my community, and maybe the world.
9. I tend to take initiative when starting new projects.
10. Wherever I have been, I have been a powerful force for constructive change.
11. No matter the odds, if I believe in something I will make it happen.
12. I am always looking for better ways to do things.
13. When I have a problem, I tackle it head-on
14. If I see someone in trouble, I help out in any way I can.
15. If I see something I don't like, I fix it.
16. I love being the champion of my ideas, even against others' opposition.

Dark Triad Definitions (the Dirty Dozen) (Jonason and Webster, 2010)

Machiavellianism

1. I tend to manipulate others to get my way.
2. I have used deceit or lied to get my way.
3. I have used flattery to get my way.
4. I tend to exploit others towards my own end

Psychopathy

1. I tend to lack remorse.
2. I tend to be unconcerned with the morality of my actions.
3. I tend to be callous and insensitive.
4. I tend to be cynical.

Narcissism

1. I tend to want others to admire me.
2. I tend to want others to pay attention to me.
3. I tend to seek prestige or status.
4. I tend to expect special favors from others.

Social Vigilantism - $\alpha=.85$ (adapted from Saucier and Webster, 2010)

1. I feel as if it is my duty to enlighten other people.
2. I feel that my ideas should be used to educate others.
3. I feel a social obligation to voice my opinion.
4. I need to win any argument about how people should live their lives.
5. Those people who are more intelligent and informed have a responsibility to educate the people around them who are less intelligent and informed.
6. I like to imagine myself in a position of authority so that I could make the important decisions around here.
7. I try to get people to listen to me, because what I have to say makes a lot of sense.
8. Some people just believe stupid things.
9. There are a lot of ignorant people in society.
10. I think that some people need to be told that their point of view is wrong.
11. If everyone saw things the way that I do, the world would be a better place.
12. It frustrates me that many people fail to consider the finer points of an issue when they take a side.
13. I often feel that other people do not base their opinions on good evidence.
14. I frequently consider writing a "letter to the editor."

The justifiability of the decision (Inman and Zeelenberg, 2002)

1. How justifiable is the decision to _____?
Weakly justifiable/Strongly justifiable
2. How easy to defend is the decision to _____?
Not easy to defend/easy to defend
3. How logical is the decision to _____?
Very illogical/very logical

Future misbehavior intentions (Daunt and Harris, 2011)

1. In the future, if it is to my advantage, I am likely to make a complaint when there is no genuine problem.
2. In the future, if it is to my benefit I may behave in a dishonest way when in a service outlet.
3. In the future, if it is to my advantage, I am likely to argue with an employee/fellow customer.
4. In the future, if I feel that it is necessary, I would be prepared to behave in a way that others within the service outlet may find unacceptable.

Intentions to Share (Lee and Ma, 2011) $\alpha = .92$

1. I intend to share news stories in social media in the future
2. I expect to share news stories contributed by other users
3. I plan to share news stories in social media regularly

Believability of Information (Sen, Gurhan-Canli, and Morwitz, 2001; Gürhan-Canli and Maheswaran, 2000) $\alpha = .92$

1. Not at all believable / highly believable
2. Not at all true / absolutely true
3. Not at all acceptable / totally acceptable
4. Not at all credible / very credible
5. Not at all trustworthy / completely trustworthy

Scenarios

Online:

Imagine that you are shopping online and see an amazing deal.

You realize that it is a pricing mistake and hope to take advantage of the opportunity before the mistake is corrected. You feel that the company might be unlikely to honor the price if you are the only customer who tries to get this deal, so you decide to let other customers know about it quickly, thinking it will be harder to say “no” to a large group of customers demanding that the company honor the deal. You realize that the best way to do that might be to share unproven or exaggerated information with the other customers to get them to request the pricing, too.

Offline:

Imagine that you are shopping offline in a traditional retail store and see an amazing deal.

You realize that it is a pricing mistake and hope to take advantage of the opportunity before the mistake is corrected. You feel that the company might be unlikely to honor the price if you are the only customer who tries to get this deal, so you decide to let other customers know about it quickly, thinking it will be harder to say “no” to a large group of customers demanding that the company honor the deal. You realize that the best way to do that might be to share unproven or exaggerated information with the other customers to get them to request the pricing, too.

APPENDIX B

ESSAY TWO

Company Description for Collection



New Latitudes is a company, which sells its brands online. New Latitudes sells everything from clothing to home goods to technology accessories. Along with selling products on its own website, New Latitudes also sells products through Amazon's marketplace, Etsy, e-bay, and others.

Figure B.1 Company Description

MANOVA Treatments

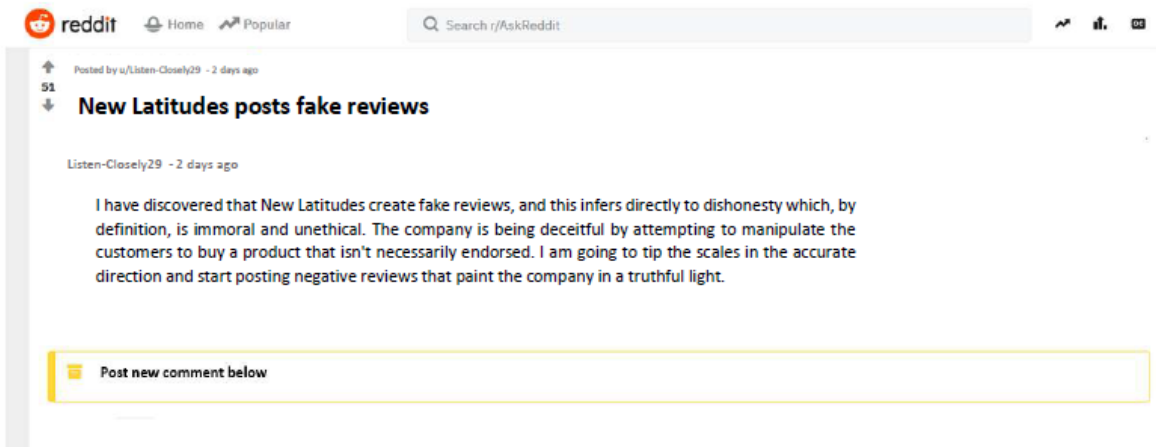


Figure B.2 Analytical Treatment

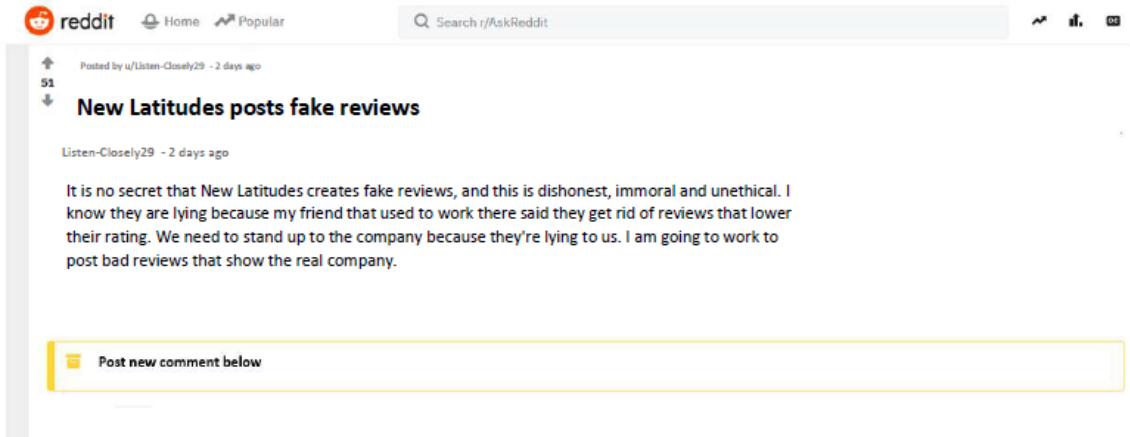


Figure B.3 Clout Treatment

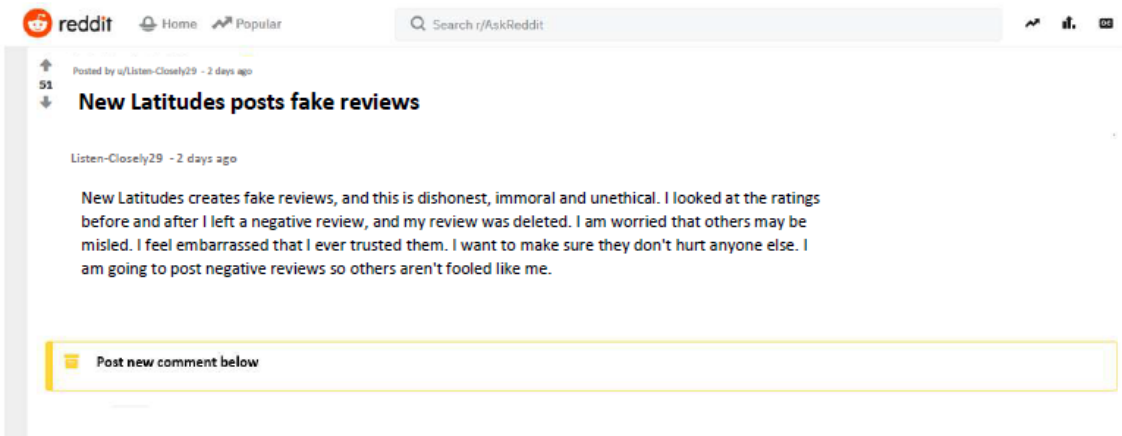


Figure B.4 Authentic Treatment

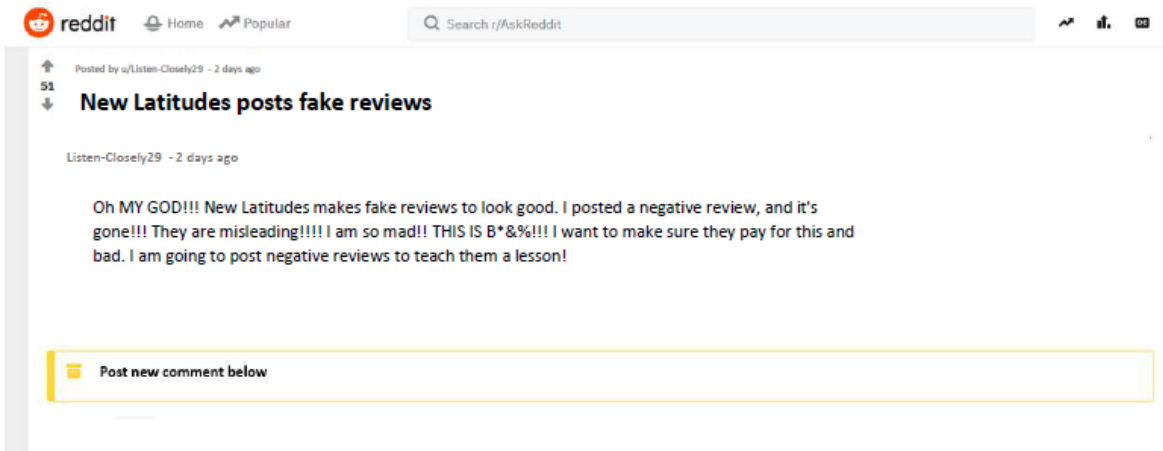


Figure B.5 Negative Emotional Tone Treatment

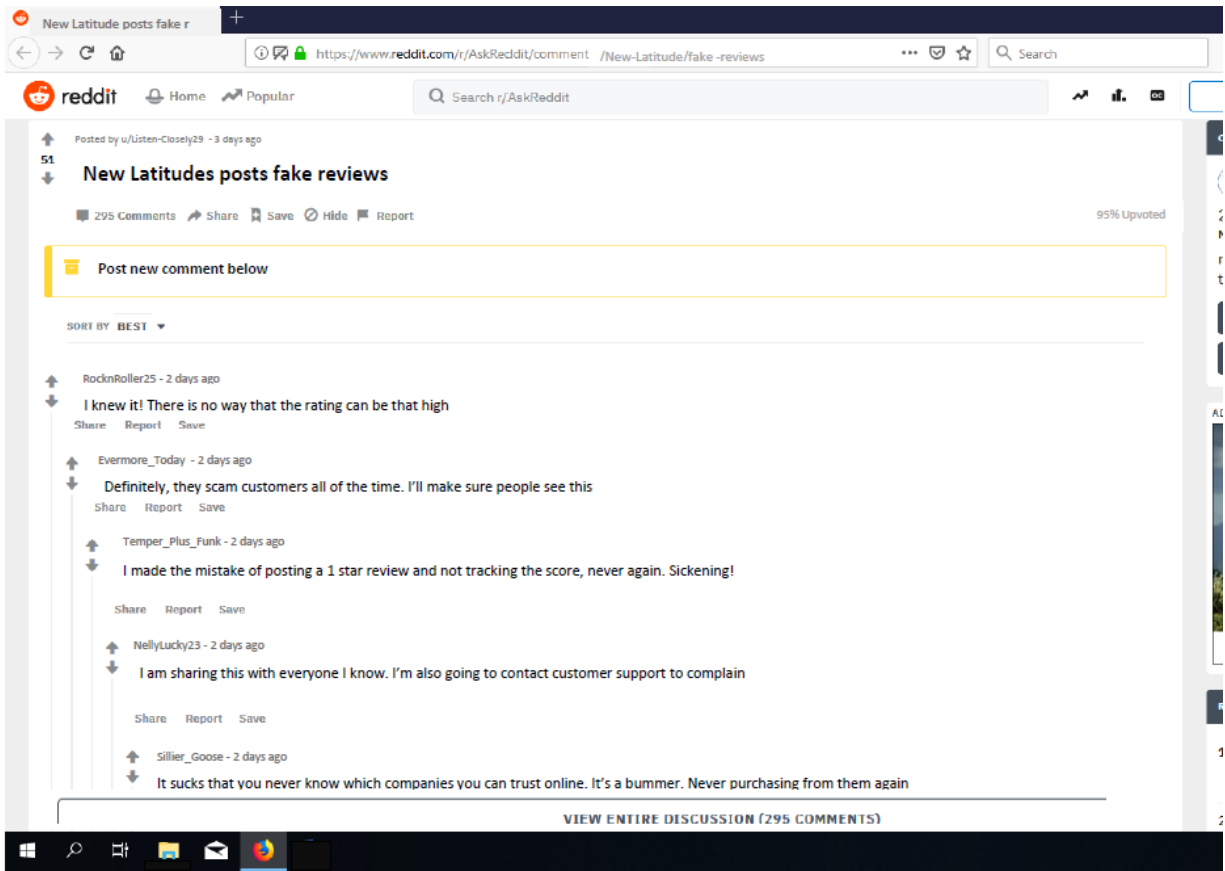


Figure B.6 Negative Cascade Treatment

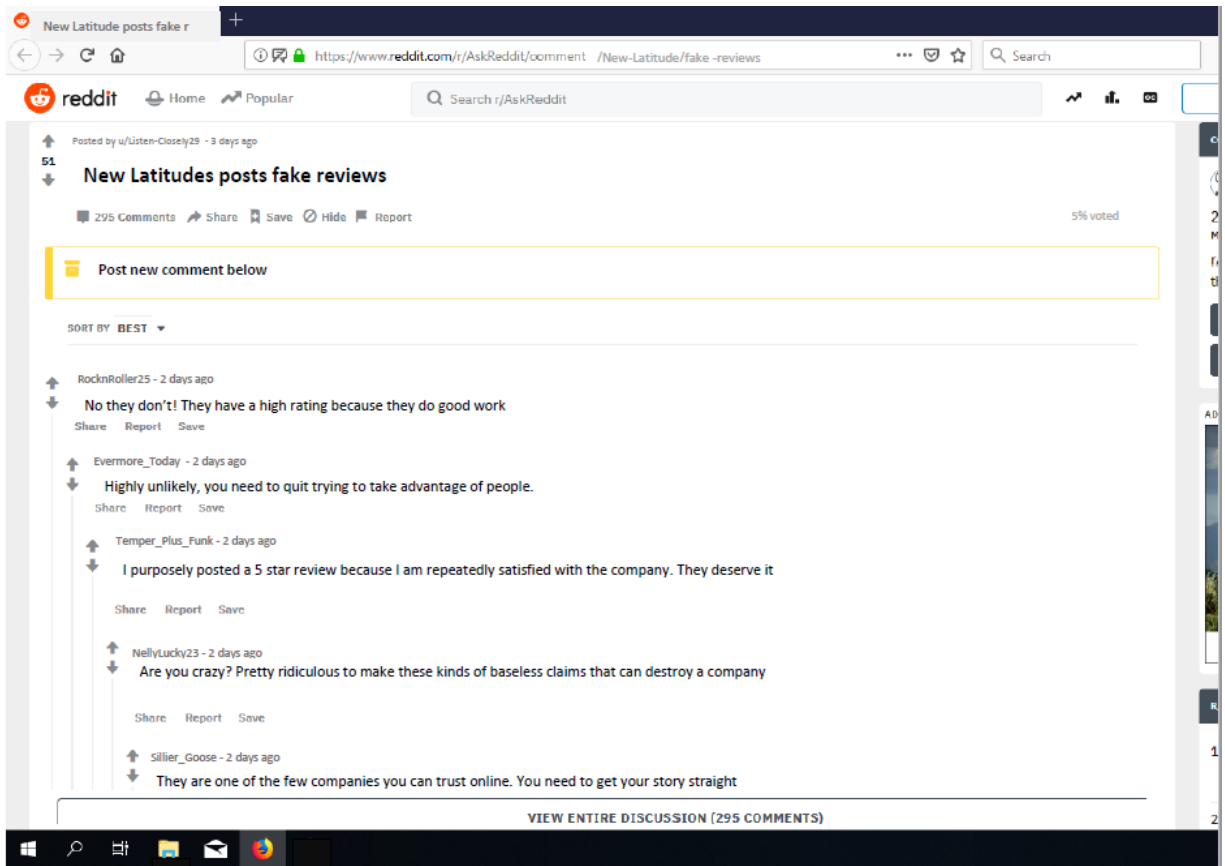


Figure B.7 Positive Cascade Treatment

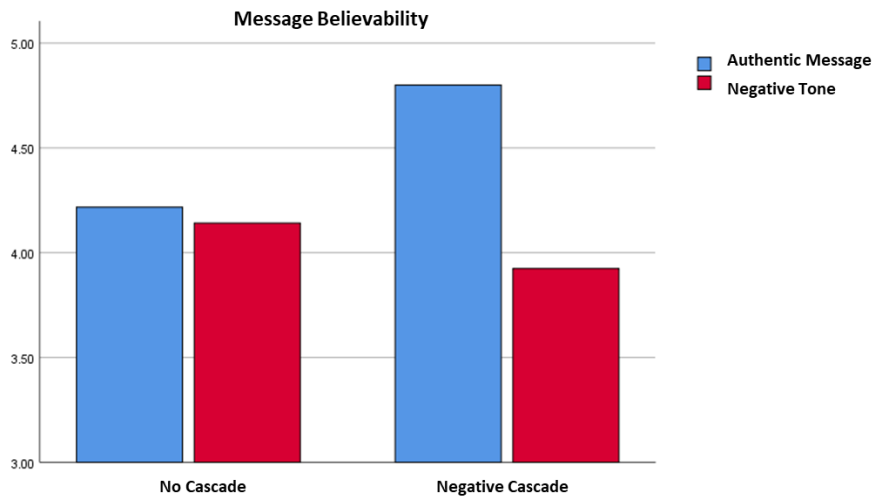


Figure B.8 Essay 2: Study 2a Interaction Visualization

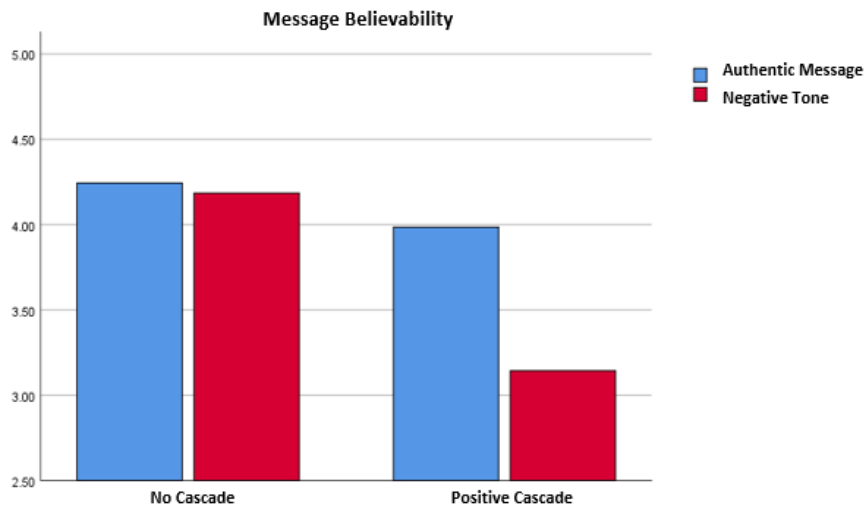


Figure B.9 Essay 2: Study 2b Interaction Visualization

APPENDIX C
ESSAY THREE

Vigilante Response Treatments

Overt Response

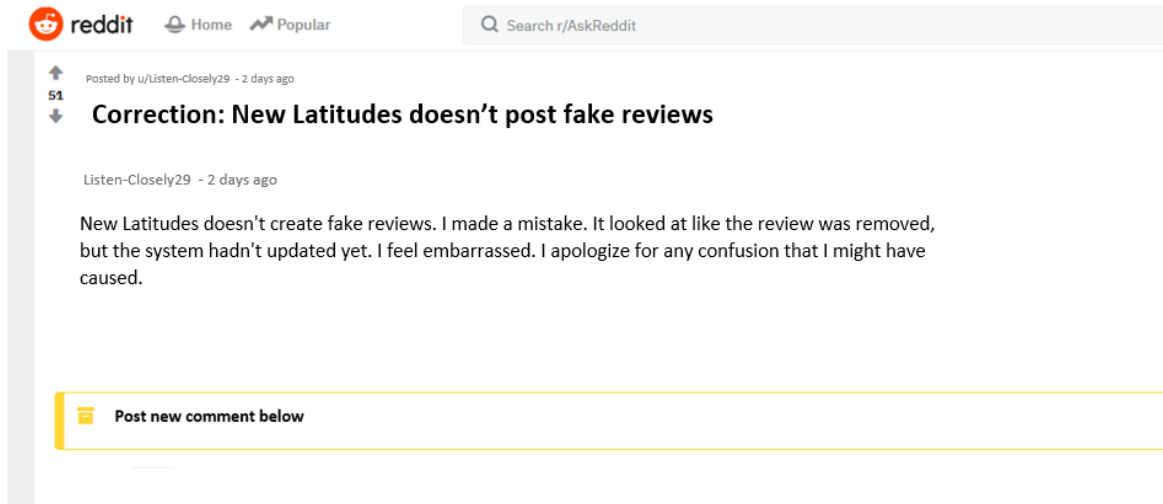


Figure C.1 Overt Response Treatment

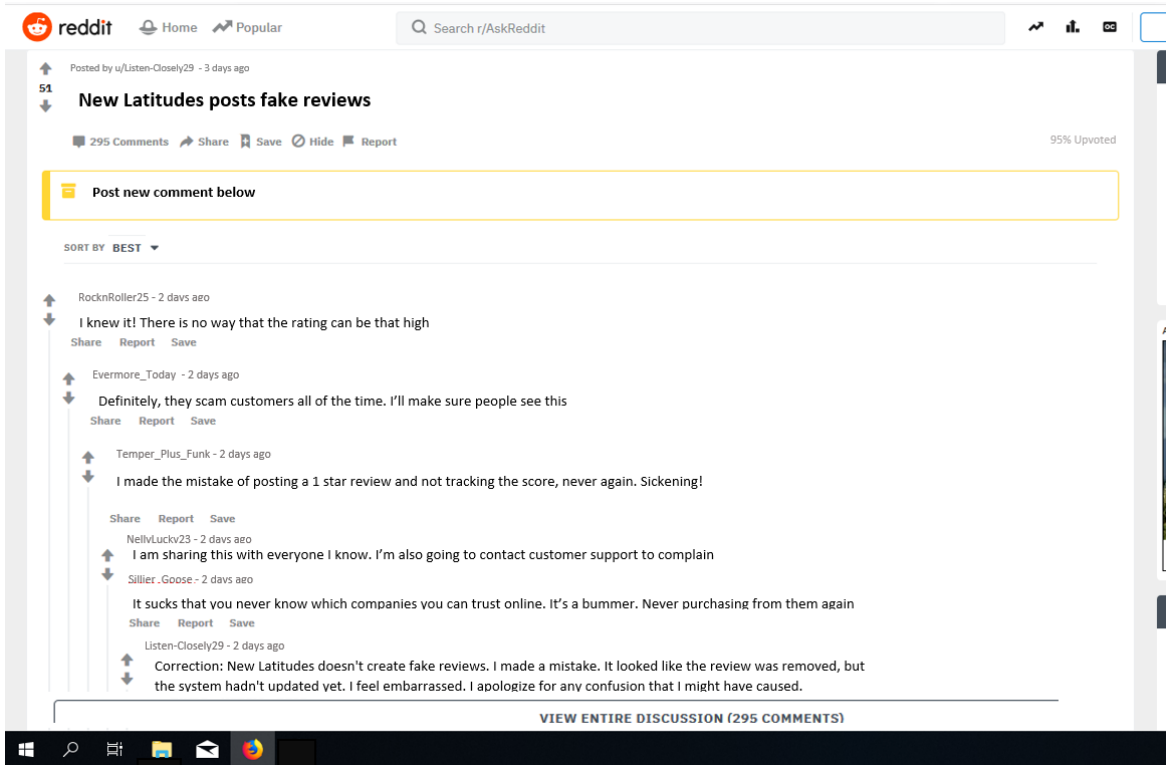


Figure C.2 Discreet Response Treatment

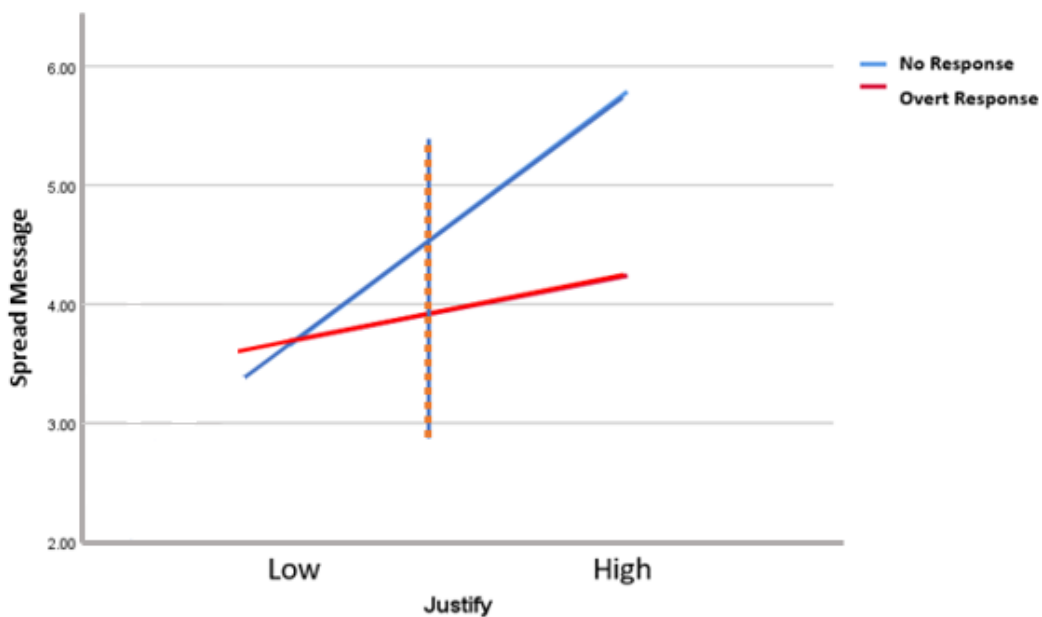


Figure C.3 Essay 3: Study 1 Interaction Visualization

Brand Quarantine Treatments

Analytical

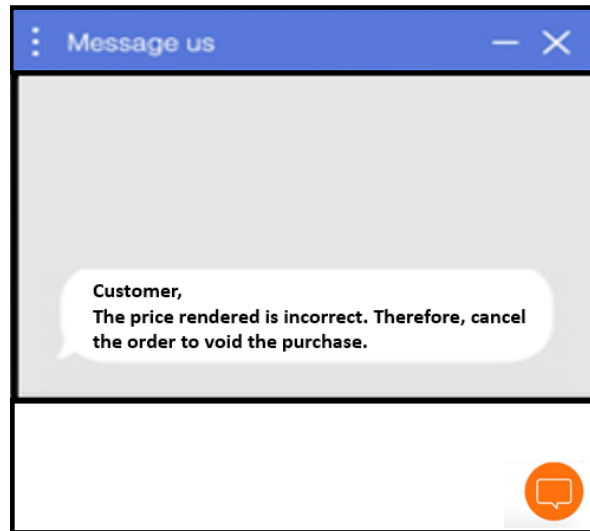


Figure C.4 Analytical Treatment

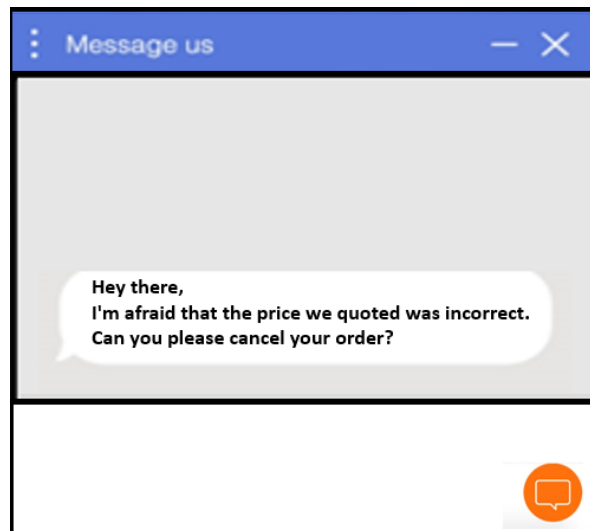


Figure C.5 Clout Treatment

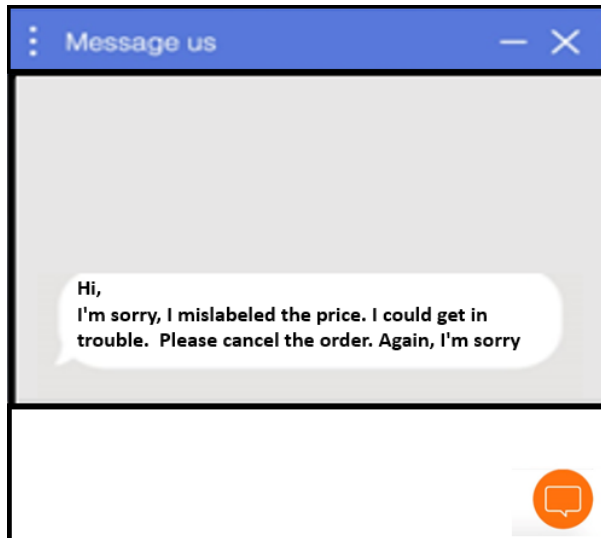


Figure C.6 Authentic Treatment



Figure C.7 Negative Emotional Tone Treatment

Essay Three: Study 2 Interaction Visualizations

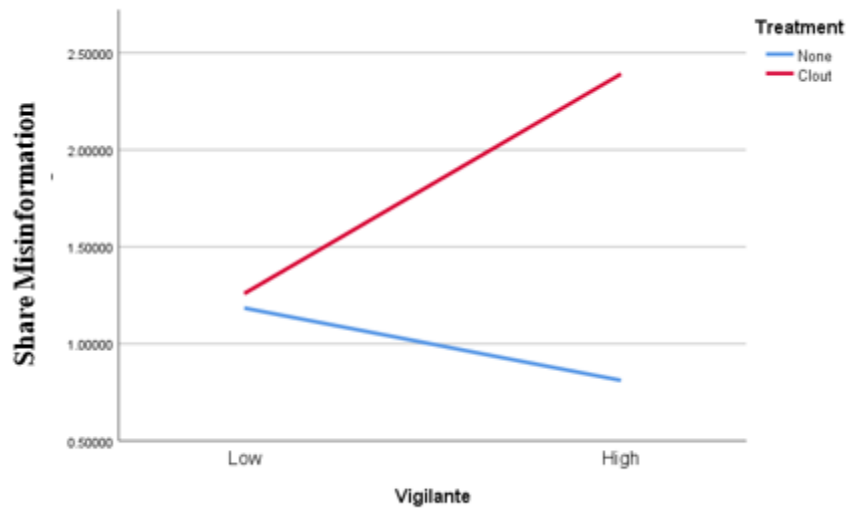


Figure C.8 Low Justification Interaction

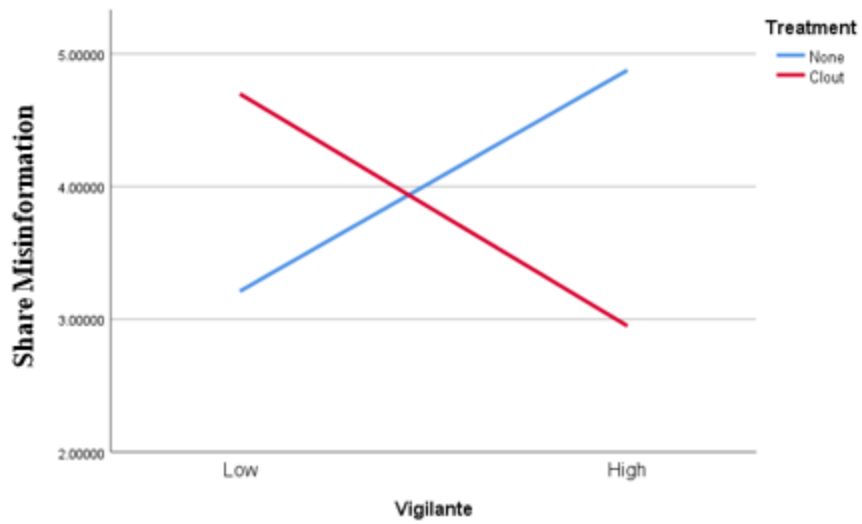


Figure C.9 High Justification Interaction

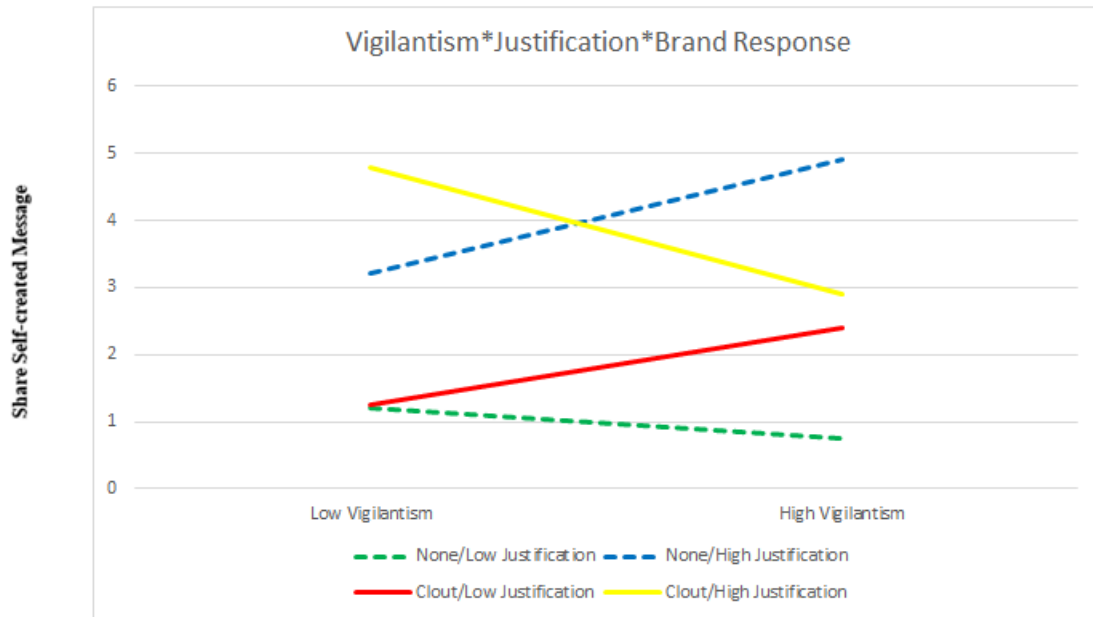


Figure C.10 Full Interaction Visualization

APPENDIX D
CONSUMER INOCULATION PROCESS

Consumer Inoculation Process and Recommended Treatments

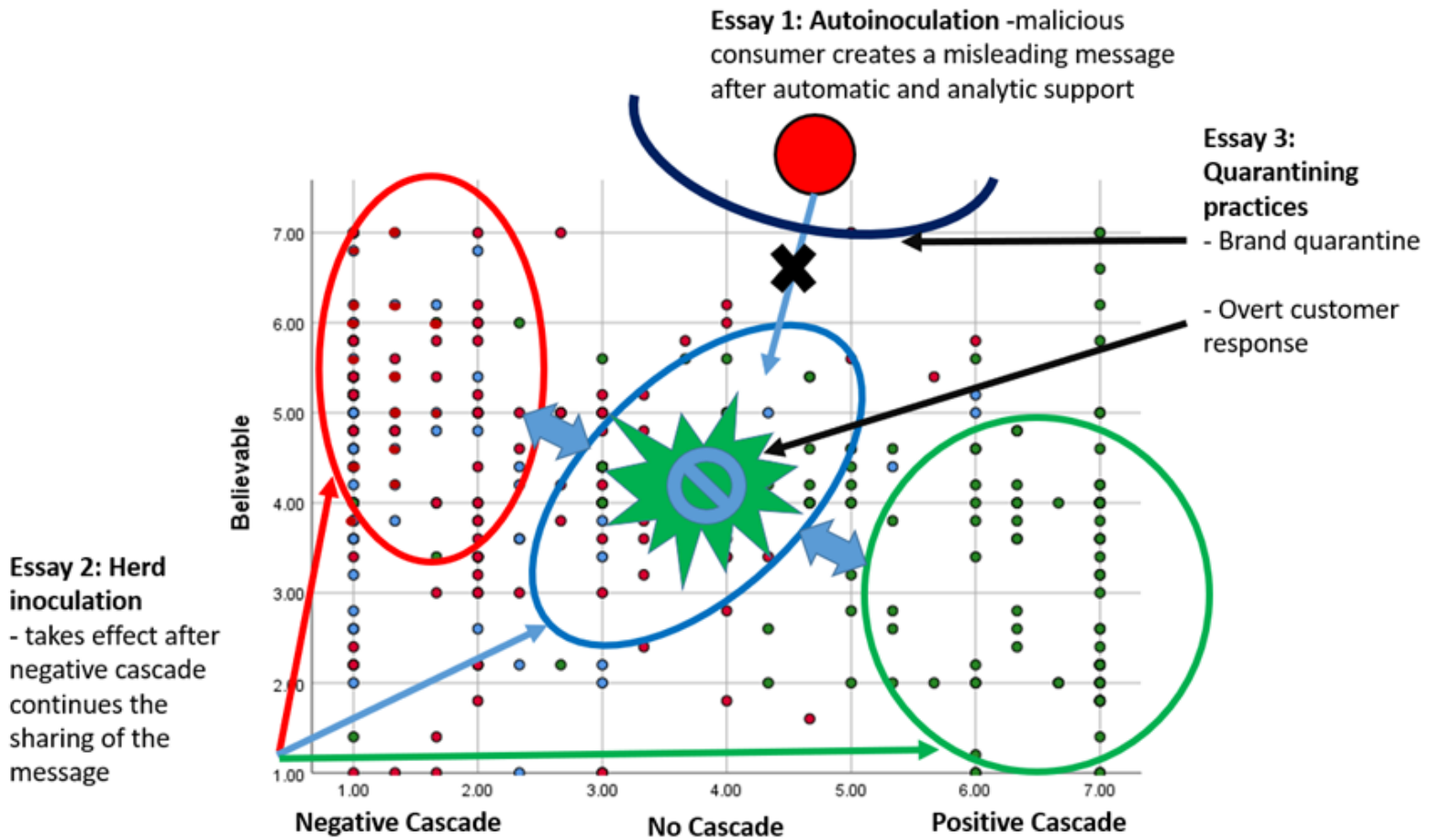


Figure D.1 Consumer Inoculation Process