

Trolling Behaviors and Victimization in Online Brand Communities

Kai Haverila

A Thesis

In the Department of

John Molson School of Business

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Business Administration)

At Concordia University

Montreal, Quebec, Canada

August 2023

© Kai Haverila, 2023

**CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Kai Haverila

Entitled: Trolling Behaviors and Victimization in Online Brand Communities

and submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (Business Administration)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____	Chair
Mahesh Sharma	
_____	External Examiner
Dr. Kai-Yu Wang	
_____	Examiner
Dr. Julie Corrigan	
_____	Examiner
Dr. Ronald Ferguson	
_____	Examiner
Dr. Marcelo Nepomuceno	
_____	Co-Supervisor
Dr. Michel Laroche	
_____	Thesis Supervisor
Dr. Michèle Paulin	

Approved by

_____ Dr. Tracy Hecht, Graduate Program Director

November 15th, 2023

_____ Dr. Anne-Marie Croteau, Dean of JMSB

Abstract

Trolling Behaviors and Victimization in Online Brand Communities

Kai Haverila, Ph.D.

Concordia University, 2023

The growth in online technology and social media use has led to a significant boom in online communication and participation. The current literature on online interactions has mainly focused on how online platforms encourage positive forms of engagement, but it is important to recognize that these platforms also create opportunities for negative types of engagement such as trolling to occur, which has become increasingly prevalent online. Currently, there is a growing academic interest in online trolling behaviors. However, the current research on trolling has some crucial limitations that must be addressed. Firstly, the trolling construct lacks conceptual clarity and trolling literature has been rather limited in scope, especially in the marketing context. To address this issue, *Essay I* conceptually explored how trolling can emerge in the brand community context. More specifically, this research introduced the brand trolling concept and developed numerous research propositions and questions that are foundational to the novel concept on the individual-, community-, and brand/organization-level. The service-dominant logic was used as a theoretical framework to illustrate the highly contextual and expansive nature of brand trolling in the brand community environment. Overall, this essay developed a more solid foundation for the trolling construct and it introduced a novel perspective on how the empirical relationships and conceptual elements of trolling can be expanded to the marketing domain on multiple levels.

Another issue prevalent in trolling literature is how its conceptualization has not been fully agreed upon by scholars and practitioners alike. Accordingly, *Essay II* addressed this issue by developing valid and reliable scales for trolling behavior and trolling victimization. Appropriate scale development procedures such as exploratory and confirmatory factor analyses, reliability tests, and numerous validity tests were conducted throughout multiple studies. The results of the studies demonstrated how trolling behavior and trolling victimization are both reliable, valid and multidimensional constructs. This research further solidifies the foundation for trolling behavior and victimization that should help scholars research the concepts more appropriately in the future.

Acknowledgements

I would like to take this opportunity to thank a lot of people who helped me along the way, because without them and their unwavering support and guidance, none of this would have been possible.

First, I want to express my sincere gratitude to my co-supervisor, Dr. Michèle Paulin, for her belief in me from the very beginning of my Ph.D. journey. Her accommodating and sympathetic approach to support and encouragement as well as her exceptional insight and flexibility throughout this experience made all of this possible. Dr. Paulin always believed in me and was willing to push me to my full potential and I will always be grateful for that.

I would also like to thank my other co-supervisor, Dr. Michel Laroche. His support and prompt feedback helped me move comfortably forward in the key stages of my Ph.D. I also want to express my appreciation to the final member of my committee, Dr. Marcelo Vinhal Nepomuceno. His insightful comments helped me gain a new perspective on my research.

I am also thankful for the faculty and support staff at the John Molson School of Business marketing department. In particular, I would like to thank Maria Dolores Mosquera, Beverley Benderoff, Tieshan Li, and Jooseop Lim, who were always willing to offer a helping hand when I needed one. They all made my experience at Concordia very enjoyable.

I would also like to thank my girlfriend, Isabelle, and my cat, Bean, for always being there for me, believing in me, and giving me some much needed love and words of encouragement after some stressful work-filled days. None of this would have been possible without your love and support.

Finally, I would also like to give the biggest thanks to my father, Matti Haverila, who was the first person to encourage me to do a Ph.D. He has always believed in me and provided me with constant support no matter what. I don't know where I would be without him. I will be forever grateful to you for giving me this opportunity.

Dedication

To my parents, Matti and Anna-Liisa. Kiitos kaikesta.

Contribution of Authors

For the first essay, the initial idea for conceptually investigating online trolling behaviors came from my term papers in two courses, Online Consumer Marketing taught by Sylvain Sénécal and Topics in Consumer Culture Theory taught by Marie-Agnès Parmentier. For both essays, I received some feedback on the overall theoretical perspective as well as the definition and conceptualization of key constructs from Michèle Paulin, Michel Laroche and Marcelo Vinhal Nepomuceno.

Table of Contents

List of Figures	x
List of Tables	xi
Dissertation Introduction	1
Essay I: A Comprehensive Conceptual Investigation of Online Trolling in a Marketing Context	3
Abstract	3
Introduction	4
Research Design	5
Theory Adaptation	5
Literature Review	5
Trolling	5
Trolling in a Marketing Context	6
Toward a Broader Perspective of Trolling	7
Brand Trolling.....	7
Theoretical Foundation	9
Social Informatics	10
Service-Dominant Logic	10
Discussion	13
Level of Analysis	13
Predictors	15
Outcomes	17
Outcomes: Value Co-Creation and Co-Destruction	18
Roles	19
Context.....	21
Implications	24
Theoretical Implications	24
Practical Implications.....	25
Directions for Future Research	26
Value Co-Creation and Co-Destruction Outcomes.....	26
Mediators and Moderators	26
Measurement of Brand Trolling.....	27
Positive Brand Trolling.....	27
Meta-Level Examination	28
Trolling by Brands	28
Individual-Directed Trolling vs. Brand-Directed Trolling	28
Management of Brand Trolling	29

Conclusions	30
Essay II: Scale Development for Trolling Behavior and Victimization.....	31
Abstract.....	31
Introduction.....	32
Literature Review	33
Trolling Behavior & Victimization.....	33
Trolling vs. Cyberbullying.....	34
Dimensions of Trolling	35
Common Method Bias	35
Item Generation and Pre-Test	36
Expert Review	37
Study 1.....	37
Sample.....	37
EFA Findings	39
Study 2.....	41
Sample.....	41
CFA Findings.....	42
Known-Group Validity	48
Predictive Validity	49
Study 3.....	51
Sample.....	51
Discriminant Validity.....	52
Validity in a Nomological Network: Trolling Behavior.....	54
Validity in a Nomological Network: Trolling Victimization	56
Concurrent Validity	59
Discussion.....	62
Results.....	62
Implications	62
Theoretical Implications	62
Practical Implications.....	63
Future Research Avenues and Limitations	64
Conclusions.....	65
Dissertation Concluding Remarks.....	66
References.....	69
Appendices.....	84
Appendix A: Summary of Research Propositions	84
Appendix B: Summary of Research Questions	86

Appendix C: Distraction Items in Trolling Behavior Scale	88
Appendix D: Deception Items in Trolling Behavior Scale	89
Appendix E: Provocation Items in Trolling Behavior Scale.....	90
Appendix F: Distraction Items in Trolling Victimization Scale	91
Appendix G: Deception Items in Trolling Victimization Scale	92
Appendix H: Provocation Items in Trolling Victimization Scale.....	93
Appendix I: List of Scales to Test Predictive Validity for Trolling Behavior	94
Appendix J: List of Scales to Test Predictive Validity for Trolling Victimization	96

List of Figures

Figure 1: Comparison of Trolling and Brand Trolling Interactions among Key Focal Actors	.8
Figure 2: Trolling Behavior Second-Order Model	46
Figure 3: Trolling Victimization Second-Order Model	46
Figure 4: Predictive Validity Model for Trolling Behavior.....	50
Figure 5: Predictive Validity Model for Trolling Victimization	51
Figure 6: SEM Model for the Nomological Network of Trolling Behavior.....	56
Figure 7: SEM Model for the Nomological Network of Trolling Victimization	58

List of Tables

Table 1: Comparing Perspectives: General Trolling and Brand Trolling Research	9
Table 2: Foundational Premises of S-D logic and Their Relationship to Brand Trolling	12
Table 3: Summary of Comparisons of General Trolling and Brand Trolling on the Micro-, Meso-, and Macro-Level.....	23
Table 4: Demographic Characteristics of Study 1 Respondents (Trolling Behavior)	38
Table 5: Demographic Characteristics of Study 1 Respondents (Trolling Victimization).....	38
Table 6: Factor Loadings for 17-item Trolling Behavior Solution (EFA).....	39
Table 7: Factor Loadings for 18-item Trolling Victimization Solution (EFA)	40
Table 8: Reliability Statistics for Trolling and Victimization Factors.....	41
Table 9: Demographic Characteristics of Study 2 Respondents (Trolling Behavior)	42
Table 10: Demographic Characteristics of Study 2 Respondents (Trolling Victimization)....	42
Table 11: Factor Loadings for 15-item Trolling Behavior Solution (CFA)	43
Table 12: Factor Loadings for 16-item Trolling Victimization Solution (CFA)	44
Table 13: Models and Fit Statistics for Trolling Behavior	45
Table 14: Models and Fit Statistics for Trolling Victimization.....	45
Table 15: CR and AVE Values of Trolling Behavior Factors	47
Table 16: CR and AVE Values of Trolling Victimization Factors.....	47
Table 17: HTMT Ratios for Trolling Behavior Factors.....	47
Table 18: HTMT Ratios for Trolling Victimization Factors	48
Table 19: Mean Scores between Heavy and Light Internet Users for Trolling Behaviors.....	49
Table 20: Mean Scores between Heavy and Light Internet Users for Trolling Victimization	49
Table 21: Demographic Characteristics of Study 3 Respondents (Trolling Behavior)	52
Table 22: Demographic Characteristics of Study 3 Respondents (Trolling Victimization)....	52
Table 23: Discriminant Validity between Cyberbullying and Trolling Behavior Factors.....	53
Table 24: Discriminant Validity between Cyberbullying and Trolling Victimization Factors	53
Table 25: Correlation Analysis for Trolling Behavior Constructs.....	56
Table 26: Correlation Analysis for Trolling Victimization Constructs	59
Table 27: Comparison between Trolling Behavior and GAIT Scales	60
Table 28: Comparison between Trolling Victimization and GAIT Victimization Scales.....	61

Dissertation Introduction

The internet has led to the creation of countless technologies and applications, which have become a ubiquitous and inseparable part of almost every individual's life and the influence of this technology is only increasing with time. Nowadays, if something needs to be done, it requires an internet connection, it is no longer an optional alternative like it was in the past. In most cases, when it comes to completing a task, the best option is to go online whether that task is paying bills, getting some work done, watching television shows, or simply communicating with family and friends. Almost everyone is connected to the online world and this gives individuals seemingly endless opportunities to instantaneously connect, communicate, and engage with like-minded individuals.

It is important to recognize that giving individuals the ability to interact with others in real-time creates opportunities for both positive and negative types of engagement to occur. One of the more prevalent forms of negative online engagement in recent times is what is known as trolling. Essentially, online trolling refers to an undirected and unprovoked behavior that mainly consists of deceptive, disruptive and deliberate actions that occur in online settings that intend to provoke reactions from random individuals (Dineva & Breitsohl, 2021; Sanfilippo et al., 2018). Trolling is also a highly relational and contextual behavior that widely differs based on the individuals and situation where it takes place. There has been a considerable amount of research interest in online trolling from numerous disciplines including but not limited to psychology (Coles & West, 2016a), information technology (Paavola et al., 2016), linguistics (Hardaker, 2010, 2013) and health science (Griffiths, 2014). However, from a marketing perspective, the research on trolling has been rather limited, which is surprising given the fact that trolling has the potential to have serious implications for brands and organizations alike (Demsar et al., 2021; Golf-Papez & Veer, 2017, 2022). Furthermore, the extant research on trolling is relatively narrow in scope where the focus is typically placed on the predictors of the behavior, its surrounding contextual elements, and how it can be managed in online communities (e.g., Buckels et al., 2014; Craker & March, 2016; Dineva & Breitsohl, 2021; Golf-Papez & Veer, 2022; Sanfilippo et al., 2018). Overall, the collective literature on trolling behavior is rather disparate and seems to lack a solid foundation. These research gaps will be properly addressed in *Essay I*, which will focus on creating a solid foundation for the trolling construct through a thorough conceptual examination of trolling from a marketing perspective.

Essay I is conceptual in nature and comprehensively investigates how trolling prevails in online marketing environments. More specifically, online brand communities are the area of focus due to their highly social and interactive nature where members can communicate and engage with others in real time, which means there is also potential for trolling to occur. Appropriately, the trolling that occurs in any marketing-related context is presented as brand trolling, which builds on the online trolling construct by taking place in online communities and provoking reactions specifically from brand community- and organization/brand-related actors. The foundation of the trolling construct is established through the introduction of several comprehensive research propositions and questions that span a wide range of topics and empirical relationships including predictors, outcomes, roles, contextual elements, and management strategies on the individual-, community-, and brand/organization-level. Given that trolling in brand communities is highly social and context-dependent, service-dominant logic is applied as a theoretical framework due to its ability to showcase how value in a community is heavily reliant on the interactions that occur among relevant actors. *Essay I* contributes to a more well-rounded comprehension of the trolling construct in the marketing domain. This is achieved through moving beyond the perspective where the main focus has

usually been the trolling individual. Instead, the focus shifts where more attention is placed on victimization as well as the roles of other actors such as the targets, bystanders, and community managers. Overall, this conceptual investigation is crucial, because not only does it introduce a novel perspective on the empirical relationships of the trolling construct in a marketing context, but it also demonstrates the expansive nature of trolling where it can impact relationships on the micro-, meso-, macro-, and even meta-level.

Clearly, it is important to introduce propositions and questions related to trolling, but it just as important to begin the process of answering them. However, this can only be accomplished if trolling has been properly conceptualized as a construct. Accordingly, conceptualization is the main focus of *Essay II*, and it is crucial because while there may be a plethora of research on trolling, there has been relatively little research on the conceptualization of trolling-related constructs (i.e., trolling behavior and victimization). This lack of research may be due to the fact that trolling is a highly contextual behavior that is often deceptive, which makes it challenging for individuals to identify the behavior when it happens. Additionally, trolling generally lacks consilience, and this is evidenced in literature where scholars create their own interpretations and measurements of the behavior (e.g., Buckels et al., 2014, 2018; Hardaker, 2010), which are often unidimensional and narrow in scope. These piecemeal conceptualizations are problematic, because they do not necessarily fully encompass the construct of trolling. This also means that there is no consensus on how trolling is defined and constructed in literature, which results in further confusion and often disjointed findings in research.

Essay II addresses these aforementioned issues through the development of a more comprehensive conceptualization of trolling behavior and trolling victimization. More specifically, this essay centers on creating valid and reliable measurement scales for trolling behavior and trolling victimization. Throughout three studies, proper scale development procedures and steps (e.g., Carpenter, 2017; Churchill, 1979; Gerbing & Anderson, 1988; Worthington & Whittaker, 2006) were conducted. Validity was rigorously assessed with tests of convergent, discriminant, predictive, known-group, nomological, and concurrent validity being completed to determine how the constructs hold up against related constructs. The results of the three studies confirm the multidimensionality, reliability, and validity of both the trolling behavior and trolling victimization constructs, where the underlying dimensions for both constructs were distraction, deception, and provocation.

Overall, the findings in both essays are significant, because they contribute to creating a more solid foundation for trolling and its related constructs, which allows for scholars to more readily conduct empirical research on trolling behavior and victimization. Together, the essays also contribute to a better understanding of trolling, which is significant, because it helps managers become better equipped to detect and manage the behavior in their respective online communities.

Essay I: A Comprehensive Conceptual Investigation of Online Trolling in a Marketing Context

Abstract

Online trolling is a construct that has been receiving increasing interest from scholars recently. However, the focus of trolling research has been quite narrow so far, especially in the marketing domain. This conceptual research introduces the concept of brand trolling, which expands the current scope of trolling research by investigating trolling from a marketing perspective. We provide a definition of brand trolling and introduce research propositions and questions that are key to brand trolling. This research will be informed by service-dominant logic, which explains the highly contextual and broadened nature of brand trolling in a brand community setting. As a construct, brand trolling moves beyond the troll-target dyad and is directed at a wide range of actors and institutions including community members, the community itself, brands, and organizations. The main purpose of this research is to generate research propositions and questions based on novel brand trolling concepts on the micro-, meso-, and macro-level. Overall, this investigation provides a more comprehensive perspective on trolling with regards to its empirical relationships and conceptual elements in a marketing context. Finally, the research concludes by discussing theoretical and managerial implications and avenues for future research with forward-looking research questions.

Keywords: Online Trolling, Customer Misbehavior, Brand Communities, Brand Trolling, Service-Dominant Logic, Social Media, Brands, Conceptual Investigation.

Introduction

Online technologies, social media platforms and online communities have become highly prominent in recent years and they have vast potential for organizations and customers alike. Online brand communities are adopted by popular brands (Baldus et al., 2015; Haverila et al., 2020) to generate mutually beneficial outcomes such as encouraging participation and engagement, enhancing customer attitudes, fostering closer relationships, and generating unique brand experiences (Baldus et al., 2015). This is favorable for customers, communities and organizations, because ideas, solutions, and brands are being discussed and shared on a large scale for everyone to view.

Online communities offer plenty of opportunities for socialization through instantaneous engagement where members can interact and discuss with one another in real-time. However, the ability to instantly interact and communicate also increases the potential for adverse behaviors such as trolling to occur. Trolling is an increasingly common behavior in online environments (Golf-Papez & Veer, 2022; YouGov, 2014), which can have widespread impacts on individuals, online communities, brands, organizations as well as broader society and its institutions (e.g., Cheng et al., 2017; Coles & West, 2016b; Demsar et al., 2021; Golf-Papez & Veer, 2017; Jane, 2015). However, despite the widespread influence that trolling can potentially have, the literature on trolling and related behaviors in a marketing context has been rather limited and disparate until recently (e.g., Breitsohl et al., 2021; Demsar et al., 2021; Golf-Papez & Veer, 2022). This reveals a research gap where the complexion of trolling is yet to be fully understood from a marketing perspective (Demsar et al., 2021).

Due to the widespread influence that trolling can potentially have on a wide range of actors and stakeholders, it becomes crucial to comprehensively understand trolling in the context of marketing. Accordingly, the primary goal of *Essay I* is to conceptually investigate trolling from a marketing perspective. The concept of brand trolling will be introduced and compared against general trolling that can occur anywhere online to determine the uniqueness of the brand trolling construct. It is also vital to properly conceptualize brand trolling to encourage empirical research on the construct, however, this will be the focus of *Essay II*. Instead, *Essay I* is more conceptual in nature and will introduce and address questions that are foundational (MacInnis, 2017) to the concept of brand trolling. Simply put, *Essay I* will pose the research propositions and questions, and *Essay II* will make it possible for them to be answered. More specifically, this research will provide an in-depth examination of brand trolling, its empirical relationships (i.e., its antecedents and consequences), its contextual elements, its relevant actors, and how actor roles have evolved over time. Furthermore, brand trolling and its related concepts will be examined on the micro-, meso-, and macro-level to assess the more widespread impact that brand trolling has on several actors and stakeholders. Finally, it is important to note that the scope of brand trolling must be bounded. The research will specifically examine the role and impact of trolls, targets, community members, and organization-related actors when brand trolling occurs within and across brand communities.

To achieve the primary goal, *Essay I* will (I) assess how brand trolling is a conceptually novel concept, (II) provide an extensive literature review that examines trolling research both in general and in the marketing context, (III) generate research propositions and questions based on the literature review, (IV) propose conceptual frameworks and tables to highlight the dynamic and multi-level nature of brand trolling, and to delineate how brand trolling is its own unique construct. The outcomes of this research are to develop a comprehensive foundation for brand trolling. Once the foundation has been set, then brand trolling can be better understood by scholars by taking the next steps of construct conceptualization and measurement. A clear understanding of brand trolling as a construct is valuable because it makes scholars better equipped to research the behavior in the future, and it also helps relevant stakeholders better

understand the behavior, which means they are better equipped to manage the behavior, which supports the long-term health of brand communities. Finally, this essay will conclude with a discussion of the theoretical and managerial implications of the findings as well as the directions for future research on the subject.

Research Design

Theory Adaptation

According to Jaakkola (2020), there are four ways to develop a conceptual paper: theory synthesis, theory adaptation, typology, and models. *Essay I* focuses on theory adaptation, which is about changing the scope of an existing theory by informing it with another perspective (Jaakkola, 2020). To put this into a trolling context, the trolling domain expands by examining it from a new theoretical lens, a marketing one specifically, where the behavior is directed at community-, organization- and brand-related actors. This further expands the scope of trolling research by switching the analysis level from dyadic relationships involving the troll to more dynamic relationships that include a wider range of actors such as trolls, targets, other community members, employees, brand communities, brands, and organizations.

The theory used to explain these new insights will mainly be service-dominant (S-D) logic, because it offers a holistic perspective that explains the more relational and contextual nature of brand trolling and how it emerges due to resources being misused in actor-to-actor interactions. S-D logic provides a good theoretical basis for expanding the scope of trolling research. Social informatics will also be used to initially inform the discussion, because it offers a solid foundation for explaining how trolling exists in the first place.

Literature Review

Trolling

The definition of trolling in this research is as follows: *trolling is an evolving, relational and multi-faceted behavior that varies in its valence, level of severity and form depending upon the actors involved and the online context in which it takes place, is intentionally performed by individuals or groups, and includes the use of deception, disruption and/or antagonism, for the purposes of provoking negative reactions from actors* (Haverila, 2023). This definition highlights the intentional, relational, and contextual nature of trolling, where the behavior can vary due to the individuals involved as well as the online setting where it takes place.

Since its origins in the 1990s (Graham, 2019), trolling is a concept that has been investigated in several disciplines, such as psychology (Coles & West, 2016a), information technology (Paavola et al., 2016), ethics (DiFranco, 2020), health science (Griffiths, 2014), linguistics (Hardaker, 2010, 2013), and law (Bishop, 2013). The literature on trolling has often focused on the predictors of the behavior, specifically related to the individuals engaging in trolling, i.e., the trolls. Research has examined predictors related to the dark tetrad of personality (i.e., narcissism, Machiavellianism, psychopathy, and sadism) (Buckels et al., 2014; Craker & March, 2016; March et al., 2017), mood and surrounding context (Cheng et al., 2017), enjoyment (Cook et al., 2017), revenge (Cook et al., 2017; Shachaf & Hara, 2010), boredom (Shachaf & Hara, 2010), and several others. However, it is noteworthy to mention that the literature has rarely considered the role of other actors or institutions in trolling situations. It would be worthwhile to examine how other actors may influence and cause trolling in order to gain a more well-rounded understanding of its predictors.

Literature has also examined the highly contextual nature of trolling. For instance, Sanfilippo et al. (2018) found that trolls are perceived differently based on individual experiences, and the community or platform where the trolling takes place. Furthermore, Cruz et al. (2018) identified the context dependency of trolling and found that trolling contains three social practices of learning, assimilating and transgressing, which are required to be successful when trolling in online communities. Finally, Cook et al. (2017) examined the contextual nature of trolling in online gaming and found that trolling can vary due to the community where it takes place and it can vary greatly in its motivations as well.

Overall, the existing literature on trolling has investigated a few subject areas in great depth, but other areas are significantly lacking. For instance, literature on trolling victimization is very scant. Hong and Cheng (2018) assessed the forms and predictors of online trolling victimization. The authors found four different types of trolling victimization, and they also discovered that depression and a sense of inferiority are significant predictors of online trolling victimization. Also, while trolling behavior has a few existing scales, trolling victimization is yet to be properly measured, which is a problem. These limited findings on trolling victimization further illustrate how the trolling construct is yet to be thoroughly studied. Accordingly, it would be highly valuable to comprehensively investigate the perspective of the targets/victims where the focus is on how trolling impacts individuals and the online communities where they take place.

Another area of trolling that has received some attention is the management of the behavior (e.g., Binns, 2012; Dineva & Breitsohl, 2021; Golf-Papez & Veer, 2022). However, most of the literature has focused on what the employees and managers can do in terms of managing the behavior. While this is an intuitive approach, it would be useful to consider what other actors such as community members could do to deal with the trolling.

Finally, another common trait of trolling research is its tendency to concentrate on the dyad between the troll and the target of the behavior, though some research has begun to examine the role of bystanders (Cook et al., 2017; Sanfilippo et al., 2018) and managers (Dineva & Breitsohl, 2021) as well. Overall, from this literature review, it becomes evident that it would be worthwhile to expand the scope of trolling literature by moving beyond the troll-target dyad and starting to focus on other relevant relationships and actors in the domain of trolling behavior.

Trolling in a Marketing Context

As mentioned in the introduction section, the literature on trolling in a marketing context has been rather limited until recently. In one of the first marketing-oriented trolling papers, Golf-Papez and Veer (2017) created a working definition and conceptual model on the manifestations of trolling behaviors. The authors also investigated how trolling could be managed and discovered that attempting to prevent the behavior may encourage it further. Finally, they discovered how opportunities for trolling can be explained by Routine Activity Theory (RAT) where crimes, or in this case, trolling is a “function of convergence of likely offenders and suitable targets in the absence of capable guardians against the offence” (Cohen & Felson, 1979).

Golf-Papez and Veer (2022) continued their research into trolling by using actor-network theory to determine how human and non-human entities allow and encourage trolling behaviors in online communities. This research explored how trolls, targets, a medium of exchange and other actors such as regulators, assistants and other trolls play a role in initiating and sustaining trolling. This draws some parallels with the current research where our goal is to specifically examine the broader role of other actors in the context of brand trolling.

Next, Nepomuceno et al. (2020) theorized that internet trolls engage in online communities to provoke users and create harmful content, consequently reducing the commercial performance (sales) in large online communities. Labrecque et al. (2022) also investigated the outcomes of trolling and found that trolling on brand posts in social media can have the unintended consequence of increasing a customer's likelihood to engage with the brand, which may be caused by a customer's need for self-affirmation (Ordabayeva et al., 2022). Ounvorawong et al. (2022) explored victimization in online brand communities and discovered how victimization (where community members are bullied by other members) negatively affected positive word-of-mouth intentions, community satisfaction and community following intentions to varying degrees. Dineva and Breitsohl (2021) adopted a unique meso-level approach to trolling by examining how organizations can manage trolling in online communities. The authors identified six distinct trolling management strategies, including direct and indirect communication approaches.

Breitsohl et al. (2018) introduced a concept similar to brand trolling known as consumer brand bullying behavior. In their research, they generated a taxonomy of consumer brand bullying behavior with six different types of behavior, which included behaviors with hostile intent (harassment, ostracism, and trolling), ambiguous intent (camouflage) and non-hostile intent (teasing and criticism). Furthermore, Breitsohl et al. (2021) delved deeper into the behavior by studying motives for engaging in customer brand-cyberbullying. The authors found that customers seeking to be popular and attractive in brand communities were more likely to engage in the behavior.

Finally, Demsar et al. (2021) contributed to a more holistic understanding of trolling by examining the social practices of trolling directed at brands and how trolling occurs in customer-brand interactions. This research opened the door for a more expansive marketing-focused perspective of trolling by introducing the concept of brand trolling and examining the impact of trolling on employees and brands.

While the literature on trolling in a marketing context has been scarce for many years, there has clearly been a recent uptick in research on the behavior. However, while there is an increasing number of papers in the domain, the findings surrounding the construct of brand trolling are still disparate and unconnected. The focus must shift from piecemeal discoveries to a more holistic examination of brand trolling that uncovers the key actors, relationships, and elements of the behavior and its victimization. Consequently, this helps build a proper foundation for the construct, which then allows for the crucial next steps such as conceptualization to be taken, which subsequently opens up opportunities for future empirical research to be conducted.

Toward a Broader Perspective of Trolling

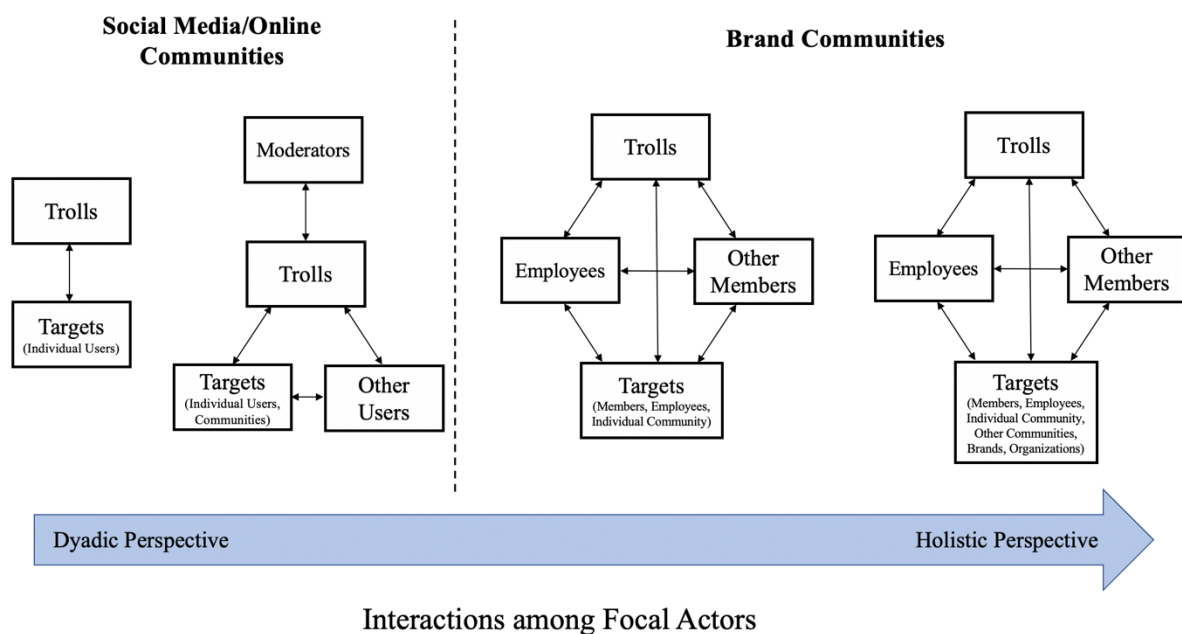
Brand Trolling

In the previous section, it became clear that trolling has been examined from a marketing perspective to a somewhat limited degree. Literature has focused on the outcomes (e.g., Labrecque et al., 2022; Nepomuceno et al., 2020; Ounvorawong et al., 2022), predictors (e.g., Breitsohl et al., 2021), management (Dineva & Breitsohl, 2021), and conceptualization (Breitsohl et al., 2018) of trolling in a marketing context. While this is a good start, more work is required to gain a fully comprehensive understanding of the behavior in the marketing domain. The starting point is to generate an appropriate definition of brand trolling. The definition will expand from the initial definition of trolling in the earlier section by considering the unique factors of brand trolling such as its focus on a wider range of actors, unique outcomes, and multi-level approach. Accordingly, the definition of brand trolling is as follows:

“Brand trolling is a type of trolling behavior that typically takes place online and is intended to denigrate anything related to the brand and provoke reactions from brand- and organization-related actors.” This definition is appropriate for all types of brand trolling that occur in a variety of different online contexts. However, for the purposes of this research, the chosen context for brand trolling will be brand communities. This setting is appropriate, because the research is specifically about trolling in the marketing context and brand communities are a highly relevant setting in marketing literature. Also, trolling in brand communities can be argued to have a more direct impact on a wide range of marketing actors including employees, brands, products, and even the organization itself. Accordingly, for this research, the definition of brand trolling will be slightly altered and is as follows: “Brand trolling is a type of trolling behavior that takes place in online communities and is intended to denigrate anything related to the brand, and provoke reactions from community-, brand- and organization-related actors.” This definition provides a solid foundation for the research propositions and questions that will be introduced later in the essay.

Compared to trolling in general, brand trolling sets itself apart in a few key ways. Brand trolling in a community context means that the relationships under consideration expand from the dyadic and triadic relationships of general trolling to a broader range of relationships that occur among a wider range of actors on the community-, brand- and organization-level. The differences between general trolling and brand trolling in terms of the involved actors and their relationships are shown in Figure 1.

Figure 1: Comparison of Trolling and Brand Trolling Interactions among Key Focal Actors



Furthermore, it is also possible to empirically examine the predictors and outcomes of brand trolling on the micro-, meso- and macro-level, which is quite extensive compared to the typical micro-level considerations that general trolling research has focused on. These differences mean that the focus of empirical research is more expansive for brand trolling compared to general trolling. Trolling research often focuses on the predictors related to trolling individuals and targets, and the outcomes related to the targets of the behavior. While this is valuable research, it is rather limited in its scope. Brand trolling in brand communities means that the research can investigate predictors and consequences related to a significantly

wider range of individuals and groups, such as targets (e.g., community members and employees), trolls, the community itself, the brand, and the organization. The key differences between trolling and brand trolling are briefly summarized in Table 1. It is important to note that it is possible for community-level factors and relationships to be explored in general trolling, however, it has not been examined yet in previous research. Accordingly, the potential aspects of general trolling (highlighted in bold font and with a *) have also been included in Table 1 to illustrate the extent to which general trolling can be investigated as a construct.

Table 1: Comparing Perspectives: General Trolling and Brand Trolling Research

	General Trolling Research	Brand Trolling Research
Focal Actors	<ul style="list-style-type: none"> ▪ Trolls, targets, bystanders, moderators*, communities* 	<ul style="list-style-type: none"> ▪ Trolls, targets, community members, employees, brand communities, brands, organizations
Context	<ul style="list-style-type: none"> ▪ Interactions occurring online (e.g., websites, communities, forums, applications, social media, gaming platforms) 	<ul style="list-style-type: none"> ▪ Interactions occurring online on brand-related communities, groups, forums, sites, applications and social media
Intended Outcomes	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of individuals, communities* 	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of communities, members, brand, etc.
Level of Aggregation	<ul style="list-style-type: none"> ▪ Micro-level, and meso-level* 	<ul style="list-style-type: none"> ▪ Micro-, meso- and macro-level
Relationships under Investigation	<ul style="list-style-type: none"> ▪ Interpersonal, dyadic, triadic, community-level* 	<ul style="list-style-type: none"> ▪ Interpersonal, dyadic, triadic, community-level (within and across), organization-level
Focus of Empirical Research	<ul style="list-style-type: none"> ▪ Predictors related to targets, trolls, and communities* ▪ Outcomes related to targets and communities* 	<ul style="list-style-type: none"> ▪ Predictors related to targets, trolls, community, and organization/brand ▪ Outcomes related to targets, brand community, and organization/brand

* = Yet to be examined in research

Theoretical Foundation

The next step is to explore existing research and use insights from it to appropriately determine the conceptual domain of brand trolling. More specifically, we will draw on existing theories and perspectives to inform the development of research propositions and questions in the arena of brand trolling.

Social Informatics

The perspective of social informatics was used as a guiding theory to explain how trolling in general emerges in online contexts. Social informatics is a perspective that examines the use of information technology in social contexts (Indiana University, 2022). More specifically, the main pillars of the perspective are that *(I)* ICT use is highly context dependent and socially situated (Sawyer & Tyworth, 2006), *(II)* people are social actors, and they have unique motivations, practices, and values that influence how and why they use ICT (Lamb & Kling, 2003; Sawyer & Tyworth, 2006), *(III)* there are social shaping of outcomes and institutions (Sanfilippo et al., 2018; Sawyer & Tyworth, 2006), and *(IV)* social outcomes are influenced by ICT use (Sanfilippo et al., 2018).

The contextual nature of ICT use explains how trolling also varies greatly due to contextual factors. Often times, trolling will be perceived differently by individuals due to their varied levels of exposure and experience with the behavior. Furthermore, online communities perceive trolling differently (Sanfilippo et al., 2018). These communities will differ in their views and perceptions due to having different expectations about acceptable behavior and conduct, a unique member base belonging to specific demographic and psychographic groups, and different types of rules, regulations, and policies in place. Accordingly, trolling will be received and perceived very differently depending upon the community where it takes place. Social informatics research also finds that ICT use can generate multiple and sometimes paradoxical effects (Sawyer & Tyworth, 2006). This occurs because online communities are sometimes quite large and contain many different types of members, resulting in multiple unique responses instead of one common one. This makes sense because not everyone will respond to a specific event in the exact same way. This is especially true in the context of trolling (Sanfilippo et al., 2018), where one group may perceive a trolling incident as light-hearted and humorous, whereas another may view it as negative and harmful.

While the social informatics perspective offers a fruitful explanation for many aspects of trolling, it mainly explains how the more general type of trolling emerges. Consequently, we must employ another perspective that more specifically informs how brand trolling emerges as a construct, and this perspective is service-dominant (S-D) logic.

Service-Dominant Logic

S-D logic is a framework that rethinks the notion of exchange and value in markets. With S-D logic, there is a shift in perspective where the focus moves from simple exchange to a view where resources and competencies can be exchanged through interactions between individuals, or actors, in a reciprocal and meaningful way (Vargo & Lusch, 2004). There is a shift from value being embedded in material products and services to value emerging from customer experiences (Prahalad & Ramaswamy, 2004). In this way, organizations are not fully responsible for creating value; instead, a wide range of actors are responsible for generating value propositions (Vargo & Lusch, 2016). With this logic, value is not only embedded in the actions and offerings of organizations, but it is also created from the actions and interactions such as sharing information, knowledge, solutions and content among community members and other organized-related stakeholders. Accordingly, community members play a vital role in generating positive and valuable brand experiences, especially in brand communities where a vast number of interactions and discussions are occurring among various actors. Furthermore, value-in-social-context (Edvardsson et al., 2011) examines how value can emerge from individual personal experiences and shared social experiences (Chandler & Vargo, 2011). The research by Edvardsson et al. (2011) suggests how an actor's perception of value in a social context depends on other customers. In the brand community context, this illustrates how value

depends on the actions of other actors and is interwoven into the interactions and socialization that occurs in the community among a wide range of unique actors.

Building on this, S-D logic also has foundational premises (FPs) that explain some of the key themes and mechanisms of the perspective. These FPs have been examined in previous literature (e.g., Vargo & Lusch, 2016) to illustrate how several unique actors can create value. However, in this research, the FPs will be used to show how value can be co-destroyed by actors engaging in brand trolling behavior. The summary of the FPs and their relation to brand trolling and co-destruction are shown in Table 2.

To start, FP6 states that “value is co-created by multiple actors” (Vargo & Lusch, 2016, p. 8), and FP8 states that “a service-centered view is inherently beneficiary-oriented and relational” (Vargo & Lusch, 2016, p. 8). These premises are important because it means that the highly relational nature of brand communities can be a unique source of competitive advantage for organizations if managed appropriately. However, when social engagement and interactions among actors become a key source of value for brand communities, this also creates a potential problem for organizations. Firstly, FP9 states that “all social and economic actors are resource integrators” (Vargo & Lusch, 2016, p. 8), which means everyone has a role to play in creating value. This premise is especially crucial in a brand community, because the lifeblood of the community are the members themselves. If there are no members, then there is no content, discussion or engagement in the community. By creating these user-centric communities, organizations give actors such as customers and community members more power by giving them the resources to freely create content and interact on a platform the organization has created. This means that content that puts the organization in a less-than-ideal light may also be created and become widespread. Secondly, while brand communities generate numerous opportunities for value co-creation through positive engagement and interactions among actors, it is also possible for the engagement and interactions to be negative, thereby creating what is known as value co-destruction. Plé and Cáceres (2010) introduced value co-destruction as an outcome that emerges “when a service system accidentally or intentionally misuses resources by acting inappropriately or unexpectedly.” This misuse of resources then leads to a decline in at least one of the systems’ well-being. This is something that organizations must begin to scrutinize because customer-to-customer interactions have been found to be a source of value co-destruction in the context of spectating sports (Kim et al., 2020; Stieler et al., 2014). From a trolling perspective, when an actor trolls in an online community, it means they are misusing the resources of being able to communicate and create content, which can negatively impact the well-being of actors including the targets of the behavior, the community, the brand and the organization. This ties into FP6 and FP9, which implied the broadened perspective of value co-creation (and co-destruction) (Vargo & Lusch, 2016). Essentially, by allowing actors such as community members to freely participate and interact in an online community, this creates a double-edged sword where actors can use the relational nature and resources of brand communities to create or destroy value (Kim et al., 2020).

Furthermore, FP10 explains how value creation (and destruction) is highly contextual and is always uniquely determined and perceived by the beneficiaries, or the recipients of service (Vargo & Lusch, 2016). This premise shows how brand trolling can be both a beneficial and harmful behavior, because individual perceptions of trolling will differ due to factors such as the context and situation (i.e., the social media platform and online community it occurred in), the contextual circumstances (i.e., how the community may have different expectations, rules and regulations with regards to acceptable behavior and conduct), and the personal experiences and roles of each actor (i.e., an individual’s tolerance and level of experience with trolling, and whether they were a troll, witness, or target) (Sanfilippo et al., 2018). Consequently, this shows how brand trolling will vary in its perception and level of acceptance

depending upon the community where it occurs. This ties into how social informatics found that ICT use can create paradoxical outcomes (Sawyer & Tyworth, 2006). Due to the increasingly prevalent use of brand communities, this leads to larger user bases, which inevitably leads to heterogenous groups emerging within these user bases. Accordingly, when a highly contextualized behavior such as brand trolling occurs, this means that different groups will perceive it in unique ways where some will view it positively whereas others will view it negatively. This shows how brand trolling can be uniquely determined and perceived by actors, thus showing how the behavior can generate value co-creation and value co-destruction depending upon the situation and actors involved.

Table 2: Foundational Premises of S-D logic and Their Relationship to Brand Trolling

Foundational Premise	Summary	Extension to Brand Trolling and Value Co-Destruction Context
FP6	<ul style="list-style-type: none"> ▪ Value is co-created by multiple actors ▪ Actors are not just the customer and organization, embraces the role of other actors 	<ul style="list-style-type: none"> ▪ Value is also co-destroyed by multiple actors, it cannot happen with one individual ▪ Actors are not just the troll and target, can also include employees, other community members, the brand, the organization
FP8	<ul style="list-style-type: none"> ▪ A service-centered view is inherently relational ▪ Actor-to-actor orientation 	<ul style="list-style-type: none"> ▪ High number of interactions and relational nature of BCs generates opportunities for brand trolling ▪ Customer-to-customer interactions can be a source of value co-destruction
FP9	<ul style="list-style-type: none"> ▪ All social and economic actors are resource integrators ▪ Implies a broadened perspective for value creation 	<ul style="list-style-type: none"> ▪ Actors can misuse resources offered to them by the BC for the purposes of trolling ▪ Implies a broadened perspective for value destruction and trolling
FP10	<ul style="list-style-type: none"> ▪ Value is always uniquely determined by the beneficiary ▪ Implies contextual nature of value creation 	<ul style="list-style-type: none"> ▪ Trolling can be perceived and determined differently due to context specificities, personal experiences and roles ▪ Implies highly contextual nature of trolling

After reviewing the FPs and their relation to value co-destruction and brand trolling, this demonstrates how S-D logic explains the more dynamic nature of brand trolling where the behavior is highly relational and contextual, more actors are included and have more involved roles, and the behavior has predictors and consequences on a broader level compared to general trolling.

As a final note, S-D logic places emphasis on the highly connected nature of actors and their interactions, which provides a solid theoretical basis for expanding the scope of trolling research with research propositions and questions. Firstly, the networked structure and

broadened perspective of S-D logic explains how the level of analysis for brand trolling can expand from micro-level considerations to meso- and macro-level considerations and even potentially meta-level factors in the future. Secondly, the broadened perspective implied in FP9 explains how brand trolling moves beyond the troll-target dyad and examines a wider range of actors including trolls, targets, brand community members, the brand community, the brand and the organization. The broader range of actors also means that the potential predictors and outcomes of the behavior are expanded, because there are more potential factors that can predict the behavior, and more potential consequences that can impact actors and institutions. Furthermore, this wider perspective also makes the management of the brand trolling more elaborate not only because there are a higher number of actors involved, but also because the connected nature of brand communities means that the impact of managing the behavior is broadened. Essentially, managing a brand trolling incident will not only affect the troll, but it may also have an impact on the community members, the community as a whole, and even the brand. Thirdly, S-D logic also addresses the highly contextual nature of the brand trolling construct. Compared to general trolling, brand trolling may be less expansive in that brand communities are only being examined, however, there are many unique contextual factors that can be investigated. For instance, brand trolling varies not only due to individual actor differences, but it may also differ due to the goals, norms, and expectations of the community as well as the product category that the brand belongs to. Clearly, the findings in this section demonstrate how S-D logic can be used to explain the intricacies of brand trolling and value co-destruction in brand communities.

Discussion

This section will specifically investigate brand trolling and its related concepts on the micro-, meso-, and macro-level. The brand trolling concepts that will be examined include *(I)* the level of analysis, *(II)* the predictors, *(III)* the outcomes, *(IV)* the role of actors, and *(V)* the context. To conclude the discussion, research propositions and questions will be developed to inform trolling research in the marketing context (see Appendices A and B for a full summary of the research propositions and questions, respectively). These propositions and questions can be utilized and further extended by trolling and marketing scholars conducting research in the domain of brand trolling.

It is important to delineate between research propositions and questions. To clarify, research propositions emerge from the previously determined findings of literature on trolling and related online behaviors and are extended into the brand trolling context. Research questions are similar in that they are inspired from previous literature findings; however, they differ because they have not been tested before, and they are more forward-looking and abstract. Research questions are also more general in nature, but they can also build off propositions and warrant a more in-depth investigation into the subject matter. For these reasons, the research questions will be included in the *Directions for Future Research* section.

Level of Analysis

Micro-Level

One of the key differentiating aspects of brand trolling is the expansive nature of the behavior compared to general trolling. When studying general trolling, attention has been typically placed on the relationships between the trolls, bystanders and targets. In this context, it also may be possible for a trolling individual to get a reprimand or punishment from moderators. However, brand trolling takes place in a brand community or platform, which

means that employees such as brand community managers will play a much more prominent role. Before we proceed, it is important to note that there are numerous online communities that are not brand-centric (e.g., hobby and/or interest-related communities), however, brand communities are the focus of this research for two reasons. First, the context is appropriate, because this essay is specifically examining trolling from a marketing perspective, and brand communities are a very pertinent setting in marketing research. Second, the community managers and employees are expected to be more formally and directly involved in the community events and discussion in comparison to the moderators of non-brand communities (e.g., subreddits, Facebook groups) who are often doing their work on a volunteer basis (Matias, 2019). The managers regularly engage through talking with community members, creating content, moderating discussions, and other related activities. Through this direct and involved engagement, it is possible for the community employees to also become targets of trolling. Accordingly, from a brand trolling perspective, this means that the relationships among trolls, bystanders, targets, and employees in a brand community would be the main areas of focus. Furthermore, this also means that predictors and outcomes related to trolls, targets, bystanders and employees would potentially be explored, which signifies how brand trolling is more comprehensive than general trolling even on the micro-level. This leads to the first section (a) of the following research proposition:

RP1a: When shifting from the general trolling construct to the brand trolling construct, the level of analysis has expanded from a triadic perspective to a multi-level perspective that includes (a) micro-, (b) meso- and (c) macro-level considerations.

Meso-Level

When the level of analysis expands to the meso-level, attention is placed on the predictors and outcomes related to the brand community itself instead of the individual actors within it. This is a vital distinction, because when brand trolling occurs, it not only impacts the individual community members and employees, but it also has an effect on the entire community. For instance, if brand trolling is commonplace in the community, it may lead to negative outcomes on the community level such as lower quality engagement and a drop in the number of community members. This broader perspective is more relevant to brand trolling, because general trolling research has tended to focus on factors and relationships on the individual-level. It is possible for meso-level elements to be examined in general online communities, however, extant research has not considered these factors in great detail yet. Nevertheless, brand communities will be the focus instead of general online communities for three reasons. First, much like in the previous section, the brand community context is more suitable for research that is investigating trolling in the marketing domain. Second, while the meso-level relationships in general online communities and brand communities might share some similarities, the effect is greater for the brand community because of the broader impact that the trolling may also have on employees, the brand, the product, and the organization. Whereas in a general online community, the impact of trolling would mainly be felt on the individual and community level. Finally, brand trolling may have a more intense effect on the community-level, because brand communities have higher expectations about trolling management compared to general online communities that often do not have dedicated managers and moderators handling them. Essentially, community members expect community managers and moderators to properly deal with trolling in a brand community. If it is not managed appropriately, then the effects on meso-level factors such as community trust and satisfaction may be quite significant. This leads to the second section (b) of the following research proposition:

RP1b: When shifting from the general trolling construct to the brand trolling construct, the level of analysis has expanded from a triadic perspective to a multi-level perspective that includes (a) micro-, (b) meso- and (c) macro-level considerations.

Macro-Level

Finally, on the macro-level, attention is placed on the antecedents and consequences of brand trolling on the brand- and the organization-level. Also, the interactions that occur within the same brand community and across different brand communities would also be examined in this context. This viewpoint is important, because it shows how brand trolling can have an impact on a broader level where it can impact organization- and brand-related elements such as brand images, reputation, and financial performance (Golf-Papez & Veer, 2017). Furthermore, this perspective can also examine the antecedents of brand trolling on a broader level. For instance, brand trolling may be directed at organizations or brands because of their controversial nature (Rotman et al., 2017). Overall, this broad perspective demonstrates how brand trolling can have causal relationships on the brand-, and organization-level, which shows the more comprehensive nature of investigating brand trolling compared to general trolling. This leads to the third section (c) of the following research proposition:

RP1c: When shifting from the general trolling construct to the brand trolling construct, the level of analysis has expanded from a triadic perspective to a multi-level perspective that includes (a) micro-, (b) meso- and (c) macro-level considerations.

Predictors

Micro-Level

On the micro-level, there are numerous individual factors that may predict brand trolling. To start, previously determined predictors of similar negative behaviors should be examined to see if they also apply to a brand trolling context. According to previous literature, factors such as sadism, psychopathy, and depression are significant predictors of trolling and cyberbullying (Buckels et al., 2014; Craker & March, 2016; Guo, 2016; Hong & Cheng, 2018). Also, agreeableness is a negative predictor of cyberbullying in previous research (Balakrishnan et al., 2019). Accordingly, examining whether these traits also apply to trolling in a brand community context would be worthwhile. Consequently, this leads to the following research propositions:

RP2: (a) Sadism, (b) psychopathy, and (c) depression positively predict brand trolling.

RP3: Agreeableness negatively predicts brand trolling.

However, it would be prudent to give attention to the more novel predictors of brand trolling. Specifically, the predictors in this section will focus on the targets of the behavior, because factors related to the targets have not received much attention in previous trolling research. Previous literature has found that individuals who have high levels of social capital online (i.e., trust and a sense of belonging) were associated with online hate victimization (Kaakinen et al., 2018). Furthermore, it was found that being an avid user of a social networking site made individuals much more likely to be a target of online hate (Costello et al., 2017). Also, social introversion has been found to be a predictor of victimization in several online and offline contexts (Festl & Quandt, 2013; Glasø et al., 2007; Olweus, 2014). For trolling

victimization specifically, Hong and Cheng (2018) found that depression was a significant and positive predictor of trolling victimization. Finally, a willingness to divulge personal information or private matters online was found to lead to victimization in individuals (Costello et al., 2017; Hawdon et al., 2014). Overall, these findings may lead to a few key conclusions. First, it shows how vulnerable individuals may be more likely victims of trolling behavior. Second, it illustrates that the more an individual shares in an online community, the more visible they are to trolls and the more likely they are to becoming a victim of trolling. This means that the more vulnerable and open members of brand communities are likelier to become targets of trolling. In a brand trolling context, the following research propositions emerge:

RP4:(a) High social capital, (b) social introversion, (c) depression, and (d) a willingness to divulge personal information positively predict brand trolling victimization.

Meso-Level

There are a number of brand community-related factors that can contribute to brand trolling. Firstly, the highly social and interactive nature of brand communities can act as an antecedent to brand trolling. Brand communities generate highly visible content, and they offer two-way communication that allows for real-time responses to any content and discussion in the community. Also, communities often make the process of creating a user profile very quick and easy, which means anyone can join the conversation almost immediately. The immediacy of interactions and the fast access to content give trolls ample opportunity to engage in brand trolling. Also, the community itself can inadvertently encourage brand trolling. If the brand community is known to be very close, then it may result in members being more likely to respond to the troll (because they want to defend the community they feel close to), which creates additional reactions, which is what trolls often desire (Cook et al., 2017). Evidently, the instantaneous nature of participation and interaction can lead to a mismatch of positive and negative outcomes in brand communities. From this discussion, the following research propositions emerge:

RP5: (a) The immediacy of brand community interactions, (b) the visibility of brand community posts, (c) customers' fast access to brand-related information, and (d) a brand community's closeness positively predict brand trolling victimization.

Macro-Level

On the macro-level, several factors related to the organization and brand can contribute to brand trolling. For instance, Rotman et al. (2017) discovered that individuals are willing to engage in unethical behaviors to punish harmful brands, and they feel no remorse over these actions. Accordingly, some individuals might engage in brand trolling to punish certain brands, which would be seen as justified from their perspective. Furthermore, the constructs of oppositional brand loyalty (Ewing et al., 2013) and inter-organization and inter-customer brand rivalry (Berendt et al., 2018) can also play a significant role in predicting brand trolling within and across brand communities due to their propensity to encourage conflict, ridicule, hostility, and other negative practices. Finally, it is possible for brands that are hated and associated with negative attributes such as moral violations and hypocrisy to be more likely targets of trolling in comparison to more trustworthy and authentic brands (Guèvremont, 2019). Similar to the findings of Rotman et al. (2017), if a brand is known to be less trustworthy and authentic, then it is likelier that individuals would punish these brands for their problematic actions. Therefore, this leads to the following research propositions:

RP6: (a) A brand's perceived harmfulness, (b) brand hypocrisy, (c) brand hate, and (d) moral violation positively predict brand trolling victimization.

RP7: (a) Brand trustworthiness, and (b) brand authenticity negatively predict brand trolling victimization.

Outcomes

Micro-Level

Numerous micro-level outcomes can emerge from brand trolling. Attention will first be placed on individual targets to see if previously determined outcomes related to victimization also apply in the brand trolling context. For instance, prior literature found that exposure to incivility leads to feelings of hostility in individuals (Gervais, 2015; Rösner et al., 2016). Similarly, trolling victimization has been found to be a significant predictor of trolling behavior (Wilson & Seigfried-Spellar, 2022), which demonstrates how negativity essentially begets more negativity. Overall, these findings show how exposure to negative behaviors can induce negative feelings and behaviors in individuals, and we believe this applies to a brand trolling context as well. Accordingly, this leads to the following research propositions:

RP8: Brand trolling victimization positively predicts (a) feelings of hostility, and (b) brand trolling behavior.

In addition, it is necessary to consider the outcomes related to the employees of the brand community, such as the moderators and brand community managers. When brand trolling occurs, it creates problems for employees in a few ways. Firstly, the resources and abilities of the employees are being misused, because they must respond to trolls and hand out punishments instead doing their main job, which is moderate content, create interesting and valuable content for the community, and respond to the questions and feedback of the community. Secondly, this increased effort that employees must give as a result of trolling incidents can potentially have a negative effect on their well-being (Demsar et al., 2021). This demonstrates how trolling in an online brand community does not only have negative consequences for the targets of the behavior and the community, but it can also have negative effects on the individuals who are responsible for managing the community itself. Accordingly, this leads to the following research propositions:

RP9: Brand trolling victimization negatively affects the (a) efficiency and (b) well-being of brand community employees.

Meso-Level

Brand trolling can also have an effect on the brand community level. When brand trolling occurs in a community, it may impact a significant number of community members which can result in negative community-wide consequences. For instance, prospective community members may be dissuaded from joining the brand community if they believe that brand trolling is commonplace in the community. Existing community members may also become frustrated with brand trolling if they feel it is becoming a normalized and expected behavior. This would have negative consequences on the quality of engagement (Breitsohl et al., 2018; Cheng et al., 2017) because individuals would be straying away from the core purpose of brand community, which is typically about content creation and constructive

discussions. If trolling persists in the community and leads to negative and lower quality engagement, this may have a ripple effect where factors such as the level of engagement, community satisfaction, community visiting and revisiting intentions, community following intentions, community trust, and positive word-of-mouth (WOM) intentions are negatively impacted by the behavior (Breitsohl et al., 2018; Dineva & Breitsohl, 2021; Ounvorawong et al., 2022). The negative reactions to brand trolling can have an impact on both the prospective and existing community members, which leads to the following research propositions:

RP10: Brand trolling victimization negatively predicts community-related outcomes such as community members' (a) joining behaviors, (b) level of engagement, (c) quality of engagement, (d) brand community satisfaction levels, (e) visiting and (f) revisiting intentions, (g) community trust levels, and (h) positive WOM intentions towards the brand community.

Macro-Level

Finally, brand trolling can also have an impact on a broader scale where the effects are felt on the brand-, and organization-level. For instance, previous literature has discovered how trolling can have a significant impact on brand reputation (Demsar et al., 2021). Furthermore, prior studies (e.g., Dineva et al., 2017; Golf-Papez & Veer, 2017) have also called for research on how trolling (and trolling management strategies) could affect customer attitudes towards the brand being trolled. It is possible that customers may view a brand less favorably if they perceive that the brand accepts trolling as a normalized behavior in their communities and social media platforms. If brands do not respond appropriately to the trolling behavior, then prospective customers may conclude that the brand is not taking proper care of their brand communities, which reflects poorly on the brand as a whole. Furthermore, if trolling continues to be a common behavior in online communities, which creates lower quality engagement, then it may have a negative impact on the community as well as the brand itself. There may be another ripple effect where factors such as positive word-of-mouth intentions toward the brand (Ounvorawong et al., 2022), brand image (Shin & Larson, 2020), brand trust (Breitsohl et al., 2018), brand satisfaction (Ounvorawong et al., 2022), and purchase intentions (Shin & Larson, 2020) are negatively affected by the persistent brand-directed trolling. This demonstrates how the effects of brand trolling are not just limited to the individuals or the community, but it can also have widespread consequences for the brand and organization. Accordingly, this leads to the following research propositions:

RP11: Brand trolling victimization negatively predicts brand-related outcomes such as (a) customer brand attitudes, (b) brand reputation, (c) positive WOM intentions towards the brand, (d) brand image, (e) brand trust, (f) brand satisfaction, and (g) brand purchase intentions.

Outcomes: Value Co-Creation and Co-Destruction

Regarding outcomes, it is also necessary to discuss how brand trolling can generate value co-creation and value co-destruction. As mentioned in the S-D logic discussion, brand trolling is able to occur in online brand communities because the organization gives members the ability and resources to create content and communicate in real-time. Consequently, this also means that individuals can misuse these resources by engaging in brand trolling, which then can lead to the decline in the well-being of at least one entity, which could include community members, the community itself, the brand, or even the organization (as seen in RP8, RP9, RP10, and RP11). Accordingly, this leads to the following research proposition:

RP12: Brand trolling contributes to value co-destruction on the individual-, community-, and organization-level.

Recent literature has also found that trolling can have positive impacts in online communities through encouraging humor, reinforcing the ethos and relationships in the community, and allowing less popular opinions to be shared, which prevents the creation of an echo-chamber (e.g., Coles & West, 2016b; Cruz et al., 2018, Hopkinson, 2013). This shows how trolling in online communities also has the potential to have a positive impact on multiple actors (Golf-Papez & Veer, 2022), which can generate value co-creation. Accordingly, this generates the following research proposition:

RP13: Brand trolling contributes to value co-creation on the individual-, community-, and organization-level.

A discussion on the more specific value co-creation and co-destruction outcomes that may emerge from brand trolling will be conducted in the *Directions for Future Research* section.

Roles

Micro-Level

A key factor differentiating brand trolling from general trolling is the broader range of actors under investigation. In this context, the targets of brand trolling are not just community members but also other brand communities, employees, the brand, and the organization. Comparatively, the roles of actors are more involved in the brand trolling context. Firstly, compared to targets of general trolling, the targets of brand trolling are better equipped to deal with the behavior. This is because most brand communities have a specific code of conduct with rules and regulations, allowing community members to report the behavior. If trolling occurred on a general online community, it may be harder for individuals to deal with the behavior. This is because many online communities either do not have a code of conduct, or their existing code is often fairly lax. Also, if trolling occurred on a general social media platform, the targets of the behavior would be much less equipped and able to deal with the behavior appropriately. This is due to the significantly larger user bases of the major social platforms, which means the management of negative behaviors is typically algorithm-based. This is a problem, because algorithm-based moderation often suffers from issues of accuracy and errors (Horne, 2023).

Secondly, the other community members who are bystanders to brand trolling incidents have heightened roles. Like targets in brand communities, they are also better equipped to deal with the behavior due to the code of conduct that gives them the ability to report the specific trolling incident. Also, when trolling occurs in a brand community instead of a general social media platform, the other community members may feel more inclined to defend the targets of trolling and respond to the trolls. This is because they have a closer relationship with the community and its members, which encourages them to protect and defend the community, because they want the brand community to flourish. These members want the community to be a central hub where sharing brand-related content, knowledge, and information is the main focus instead of enforcing rules and reporting bad behavior. Furthermore, previous research found that brand defense emerges because individuals who identify with a brand (such as brand fans) view a threat to the brand as a threat to the self (Lisjak et al., 2012). This supportive

behavior has been noted in previous research where defending online communities from trolling contributed to community building, and reinforcing relationships among the community members (Hopkinson, 2013). Overall, there is considerable amount of research (e.g., Hassan & Casaló Ariño, 2016; Japutra et al., 2014; Lisjak et al., 2012) on individuals engaging in brand defending, which demonstrates the more involved role of the “other community member”. However, one could argue that this type of defending behavior is much less likely to happen on a general social media platform where the relationships are typically more distant, and users have little reason to engage in supportive behavior.

Thirdly, brand trolling is unique because the trolls can direct their behavior at the community, employees, the brand, or even the organization. Furthermore, brand trolling can occur within the community among community members or across different brand communities. It is also possible for multiple actors to engage in brand trolling such as the troll, targets, bystanders, or even the brand itself. Finally, trolls have higher expectations in the sense that they must possess a certain level of knowledge about the community and brand to succeed in their trolling; they must assimilate into the community (Cruz et al., 2018). This would not be as necessary for general trolling because the knowledge required to interact on a popular social media platform or general online community is often much less specific and involved compared to a brand community. Overall, this demonstrates how trolling in a brand community is a more involved effort compared to general trolling in online environments.

Finally, the employees of the brand community must manage trolling in a more involved manner where they may have to provide punishments, issue warnings, reply to the involved parties (the troll and the target), or act as a moderator and regulator. For employees, these tasks go beyond their expected role that consists of creating brand-related content and replying to member questions, comments and feedback. The expectations for brand community employees are higher than a general social media manager when dealing with trolling behavior. The reason for the higher expectations is that brand communities have a specific code of conduct that must be followed, and they also have a dedicated group of moderators whose job is to enforce the rules laid out in the code of conduct. This must be done so that the brand community can protect itself from negative behaviors and remain as a place where focused brand-related content, ideas, and discussion can be shared among devoted community members. Also, the relatively low number of members in a brand community makes the moderators' jobs a realistic task in this setting, which means members expect the behavior to be properly dealt with. In contrast, moderation on a site-wide scale is nearly impossible on general social media platforms due to their extensive user bases. While social media platforms also have their own respective codes of conduct, the moderation and enforcement of these codes is usually done by algorithms, which have some accuracy limitations (Horne, 2023). Furthermore, the expectations of brand community employees can be argued to be higher compared to moderators of general online communities. Essentially, because of their professional nature and expectation of direct action and communication, brand community employees are expected to swiftly and properly deal with trolling in brand communities. However, the expectations for moderators of non-brand communities are relatively lower, because they are often doing this work on a volunteer basis (Matias, 2019), which means they cannot dedicate the same amount of time and effort as a salaried employee to deal with trolling.

It is also vital for employees to know how to manage and respond to trolling in the community. As previously mentioned, trolling can potentially have positive outcomes for the community as well, which raises the question: should trolling be allowed in brand communities, and if so, to what extent can brand community employees allow for brand trolling to exist? This question is all about setting the right expectations for the community. You do not want to have a purely zero-tolerance approach, because the response may come across as overly harsh and a violation of freedom of expression to some users. Consequently, this type of response

could negatively impact factors such as customer attitudes toward the community and brand, the level of community engagement, the satisfaction and trust levels toward the community and brand, and so on. However, you also do not want to have a fully relaxed *laissez-faire* approach, because that can lead to the brand community becoming a so-called “wild west” where trolling is normalized, and the content and discussion are unfocused and unrelated to the intended purpose of the brand community. Overall, this shows how the organization must strike a healthy balance between protecting the community and its members and stimulating and allowing diverse opinions and expressions. Accordingly, employees must learn to manage brand trolling appropriately to mitigate its negative effects and also be conscious of how their management approach may impact the community and brand as a whole. This creates an interesting crossroads for the brand community employees in terms of their overall role and expectations. Accordingly, the lengthy discussion in this section leads to the following research propositions:

RP14: When shifting from the general trolling construct to the brand trolling construct, the roles and expectations have significantly changed for (a) the targets of trolling, (b) the community members (c) the trolls themselves, and (d) the employees.

Context

Meso-Level

The contextual and situational elements of brand trolling differ significantly from that of general trolling. Firstly, the overall context is narrower for brand trolling, because the behavior is situated mainly in brand communities and brand-related sites, applications and platforms. This is a much narrower context compared to general trolling, which can happen anywhere online including comment sections, websites, applications, forums, social media platforms, online gaming platforms and so on (Duggan, 2017; Golf-Papez & Veer, 2022). Secondly, as mentioned in the previous section, there are a wider range of actors that play a more prominent role in brand trolling compared to general trolling.

Most importantly, there are many unique factors of brand communities that may encourage or discourage brand trolling. Firstly, brand communities can differ due to their goals, norms, expectations, content, and many other factors. This is possible, because the community employees may have different ideas on how to do their tasks such as create content and enforce the rules of the brand community. Some employees may allow trolling behavior due to the brand community’s established norms and expectations regarding what is and is not acceptable behavior. Other employees may simply be more lax with regards to the rules and how they are enforced. Finally, some communities may have more experience and exposure to trolling, which means the employees know how to appropriately manage and respond to the behavior, and they may even partially accept it knowing that it is nearly impossible to fully suppress the behavior.

Another way in which brand communities can differ are their overall goals and type. Relling et al. (2016) investigated how the type of community would impact positive and negative word-of-mouth intentions in online brand communities. The authors found that negative word-of-mouth caused more negative reactions in social goal communities compared to functional goal communities. In a similar vein, it would be intriguing to determine if the incidence of brand trolling would vary based on goals that are specific to brand communities. For instance, brand communities may differ due to focusing on different goals related to performance (e.g., increasing sales, offering discounts, codes and perks to members), information (e.g., offering newsletters, and informative content), or engagement (e.g.,

encouraging likes, sharing, comments and reactions). In another study, Jung et al. (2014) found that unofficial brand communities were not as effective as official brand communities in enhancing brand trust. Accordingly, it would be interesting to investigate how brand trolling and its potential impact may be affected by whether the behavior takes place in a marketer-created official brand community versus a fan-created unofficial brand community.

Finally, the incidence of brand trolling could also vary based on more general community factors such as its “sense of community” (or closeness), and size. It seems intuitive that trolling would be more likely to occur in brand communities where the behavior would receive more attention through having more potential targets, more posts to comment on, and a higher likelihood of response to the behavior. Consequently, these differences in brand communities leads to the following research propositions:

RP15: Brand trolling differs in its frequency, severity, and manifestation across online communities with (a) different norms, rules, and expectations, (b) different levels of experience with the behavior, (c) different goals, (d) varying levels of “officialness”, (e) closeness, and (f) community size.

Macro-Level

The contextual elements of brand trolling are also unique on the brand- and organization-level. For instance, brand trolling may be more likely to occur for products and brands with larger user bases. This is because there are more potential targets for trolls, which would increase the likelihood of the behavior gaining attention. Furthermore, it is possible for brand trolling to differ based on the type of product that the brand and brand community belong to. For instance, previous literature (Ounvorawong, 2021) has called for research that investigates the differences between brand victimization for individuals who support high-involvement products (e.g., clothing) and low-involvement products (e.g., fast food). It is possible that brand communities of high-involvement products may have a higher propensity to being trolled, because the community members have more of an emotional attachment to high-involvement products in comparison to low-involvement products. Furthermore, brands with high self-relevance were more likely to cause brand-related aggression than brands with lower self-relevance levels (Johnson et al., 2011). In a brand trolling context, this could mean that trolling is more likely to be directed at brands that people feel are a part of their self-concept or self-image, because when someone is trolling that brand, then they are also trolling the person who enjoys and consumes the brand, which falls in line with the findings of Lisjak et al. (2012). This demonstrates how there is potential for brand trolling to be influenced by the type of product and the level of importance that the product/brand holds for individuals. This warrants further investigation and leads to the following research proposition:

RP16: Brand trolling differs in its frequency, severity, and manifestation across online communities of brands (a) with varying user base sizes, and with varying levels of (b) involvement and (c) self-relevance.

Overall, the discussion on the main concepts and research topics for brand trolling on the micro-, meso- and macro-level, and how it compares to general trolling is fully summarized in Table 3 (see Appendix A for a full summary of the research propositions).

Table 3: Summary of Comparisons of General Trolling and Brand Trolling on the Micro-, Meso-, and Macro-Level

	General Trolling (Micro-, and Meso-Level)	Brand Trolling (Micro-Level)	Brand Trolling (Meso-Level)	Brand Trolling (Macro-Level)
Key Focus	<ul style="list-style-type: none"> ▪ Micro-level relationships of trolls, targets, and other users ▪ Meso-level relationships of communities 	<ul style="list-style-type: none"> ▪ Micro-level relationships of trolls, targets, other members, and employees 	<ul style="list-style-type: none"> ▪ Meso-level relationships of brand communities 	<ul style="list-style-type: none"> ▪ Macro-level relationships of brands and organizations
Focal Actors/Entities	<ul style="list-style-type: none"> ▪ Troll ▪ Target ▪ Other users ▪ Moderators ▪ Communities 	<ul style="list-style-type: none"> ▪ Troll ▪ Target ▪ Community members ▪ Employees 	<ul style="list-style-type: none"> ▪ Brand communities 	<ul style="list-style-type: none"> ▪ Brand communities ▪ Brands ▪ Organizations
Roles	<ul style="list-style-type: none"> ▪ Can change and evolve ▪ Actors are less involved ▪ Actors are not well equipped to deal with behavior 	<ul style="list-style-type: none"> ▪ Can change and evolve ▪ Actors are more involved, and more effort is required from them ▪ Actors are ready to deal with behavior 	-	-
Context	<ul style="list-style-type: none"> ▪ Online communities and groups ▪ Websites ▪ Forums ▪ Social media ▪ Applications ▪ Gaming platforms 	<ul style="list-style-type: none"> ▪ Online brand communities and groups ▪ Brand-related sites, forums, social media, applications 	<ul style="list-style-type: none"> ▪ Online brand communities and groups ▪ Brand-related sites, forums, social media, applications 	<ul style="list-style-type: none"> ▪ Online brand communities and groups ▪ Brand-related sites, forums, social media, applications
Predictors	<ul style="list-style-type: none"> ▪ Personality and psychological factors of troll and targets ▪ Demographics of trolls and targets ▪ Factors related to community content and its closeness 	<ul style="list-style-type: none"> ▪ Personality and psychological factors of trolls and targets ▪ Demographics of trolls and targets 	<ul style="list-style-type: none"> ▪ Factors related to brand community content ▪ Factors related to closeness of brand community 	<ul style="list-style-type: none"> ▪ Factors related to positive and negative aspects of organization/brand

	General Trolling (Micro-, and Meso-Level)	Brand Trolling (Micro-Level)	Brand Trolling (Meso-Level)	Brand Trolling (Macro-Level)
Potential Outcomes	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of targeted individual and community 	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of targets, members, and employees 	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of community 	<ul style="list-style-type: none"> ▪ Provoking reactions, disruption, deception ▪ Denigration of organization/brand
Relationships under investigation	<ul style="list-style-type: none"> ▪ Interpersonal ▪ Dyadic ▪ Triadic ▪ Community-level 	<ul style="list-style-type: none"> ▪ Interpersonal ▪ Dyadic ▪ Triadic 	<ul style="list-style-type: none"> ▪ Community-level (Within Community) 	<ul style="list-style-type: none"> ▪ Community-level (Within and Across Communities) ▪ Organization-level/Brand-level ▪ Ecosystem-level
Focus of empirical research	<ul style="list-style-type: none"> ▪ Predictors related to targets, trolls and community ▪ Outcomes related to targets and community ▪ Contextual elements related to community 	<ul style="list-style-type: none"> ▪ Predictors related to targets and trolls ▪ Outcomes related to targets and employees ▪ Expanded roles of individual actors 	<ul style="list-style-type: none"> ▪ Predictors related to community ▪ Outcomes related to community ▪ Contextual elements related to community 	<ul style="list-style-type: none"> ▪ Predictors related to organization/brand ▪ Outcomes related to organization/brand ▪ Contextual elements related to organization/brand

Implications

Theoretical Implications

The goal of this research was to conceptually investigate trolling from a marketing perspective and introduce the concept of brand trolling. More specifically, the aim was to expand the empirical understanding of the trolling concept beyond its often researched dyadic and micro-level outlook and consider factors on the broader meso- and macro-level. This was accomplished by shifting to a more holistic perspective that included examining a wider range of actors where their roles, expectations and relationships are much more involved. For instance, in a brand community context, the actors are more inclined to appropriately deal with the behavior due to their more involved role and the sense of community that has developed in the community. Furthermore, the relationships under examination move beyond the individual level, and consider the broader impact of the behavior on the community and organization. This demonstrates how brand trolling does not occur in a vacuum, the behavior will impact numerous actors and institutions on a wider scale, and this must be further investigated in the future.

Next, this conceptual research introduced the novel empirical relationships that may surround the brand trolling construct. More specifically, using previously determined predictors and consequences from prior research and extending them into the brand trolling context was beneficial, because investigating these relationships and how they can impact a wide range of actors in the novel brand community context can provide some valuable insights about brand trolling and online behaviors in general. Overall, this research demonstrates how brand trolling can go into greater depth and detail compared to general trolling, because the empirical relationships move beyond predictors and outcomes related to the trolls and targets. The more exhaustive nature of brand trolling is evident when the empirical relationships can focus on predictors and outcomes related to trolls, targets, employees, the community, and the brand. This illustrates the more in-depth nature of trolling that occurs in a marketing context.

Finally, brand trolling seems to be more expansive in terms of its contextual elements. The boundaries of brand trolling are narrower compared to general trolling in the sense that brand communities and groups are the main areas of focus instead of websites, applications, social media platforms, gaming platforms and so on (Duggan, 2017; Golf-Papez & Veer, 2022). However, with regards to unique situational factors, brand trolling can be viewed from a multitude of different perspectives due to the great diversity that exists among brand communities. For instance, brand trolling can vary due to the specific goals, norms, and rules of the brand communities whereas general trolling taking place on large social media platforms would be more challenging to examine, because the overall content and discussion that occurs on the platforms is quite broad and often lacks specificity. Overall, this shows how more depth can be brought into the investigation of brand trolling compared to general trolling, because the frequency, severity and manifestation of brand trolling can differ due to a variety of factors that go beyond the elements related to the troll or target of the behavior.

Practical Implications

There are some implications that emerge from a practical perspective as well. Introducing the concept of brand trolling suggests the existence of trolling in online brand communities, which makes the communities readier and better equipped to manage the behavior. Accordingly, communities can introduce more relevant rules and regulations, which set the right expectations about the expected behavior and conduct in the community. This alone will not fully eradicate trolling, but it is a good start to minimize the likelihood of it occurring in brand communities.

When managing trolling, managers must be aware of the potential ramifications that may emerge from these strategies. When responding to the behavior, managers must strike an appropriate balance between a fully zero-tolerance and *laissez-faire* approach. Managers should not adopt a zero-tolerance approach, because the brand community may be viewed as a place where members cannot express their honest opinions and humor is not allowed. However, they should not adopt a *laissez-faire* approach either, because this may lead to trolling becoming normalized, which may consequently lead to senseless and inappropriate content and discussion becoming a norm in the community. This is a key decision for managers, because the response to brand trolling may affect community- and organization-related outcomes such as joining behaviors, customer attitudes, brand/community reputation, brand/community satisfaction, and many others. This demonstrates how decision-makers must strike a healthy balance when managing brand trolling in a community. Managers must set the right expectations so that the brand community can serve its key purpose of being a hub where devoted community members can openly interact as well as create and share brand-related content.

Finally, in the brand trolling context, the role of bystanders and other community members is heightened. Managers should be aware of this and encourage them to act in the appropriate manner when brand trolling occurs. This is crucial, because community members are equipped to report trolling behavior in a brand community, and they also may be more inclined to help if they feel close to the community. Knowing this, managers should foster a close, tight-knit brand community to encourage the members to speak up and defend the community when trolling occurs. This not only creates a sense of community, but it also makes the community members feel like they are appreciated and being listened to.

Directions for Future Research

While there were several research propositions related to brand trolling that were introduced in this research, there are some additional and more specific research avenues and questions that emerge from the research propositions, which should be explored in the future. The following section will discuss these emerging research questions (see Appendix B for a full summary of the research questions).

Value Co-Creation and Co-Destruction Outcomes

The earlier discussion that prompted RP12 and RP13 introduced how brand trolling can potentially contribute to value co-creation and co-destruction in general. While this is valuable, it would also be worthwhile to investigate the more specific outcomes of value co-creation and value co-destruction in the brand trolling context. Previous research has generated several dimensions of value co-creation such as collaboration, helping and information sharing (Busser & Shulga, 2018; Yi & Gong, 2013). Also, some underlying dimensions of value co-destruction have been generated such as poor interpersonal communication and irresponsible behavior (Guan et al., 2020). However, it is important to note that the studies that generated these dimensions (i.e., Busser & Shulga, 2018; Guan et al., 2020) were examined in the context of tourism and hospitality. It would be interesting to determine how the findings would differ when studied from a brand community and brand trolling perspective. More specifically, it would also be worthwhile to assess what specific value co-creation and value co-destruction outcomes emerge in brand communities as a result of brand trolling. Furthermore, examining what specific type of value (e.g., social, emotional, hedonic) would be impacted by brand trolling would also be an interesting avenue of research. From this, we postulate the following research questions:

RQ1: What specific value co-creation outcomes emerge from brand trolling in online brand communities?

RQ2: What specific value co-destruction outcomes emerge from brand trolling in online brand communities?

Mediators and Moderators

In empirical research, there are numerous factors that could act as mediators and moderators in relationships between constructs, and brand trolling is no exception. In prior trolling literature, factors such as loneliness (Masui, 2019) and moral disengagement (Wu et al., 2021) have been found to be moderators while social media fatigue (Wu et al., 2021), and trolling enjoyment (Buckels et al., 2014) have been found to be mediators in trolling frameworks. Prior literature has called for further research on potential mediators and moderators in trolling such as anger, depression, empathy and moral identity (Wu et al., 2021).

Knowing this, it would be worthwhile to see if there are any potential mediators and moderators that are specific to brand trolling. For instance, could the length of brand community membership or the size of the brand community play a potential mediating or moderating role in some empirical relationships? Accordingly, we postulate the following research question:

RQ3: What constructs could potentially act as moderators and/or mediators in a brand trolling framework?

Measurement of Brand Trolling

Another potential area of future research is the measurement of brand trolling. Currently, there are useful tools such as Perspective API that can be used to identify toxicity in comments. However, it would also be valuable to come up with a measurement scale of brand trolling that could be used in empirical research. Investigating the conceptualization of brand trolling and discovering the specific dimensions and items that make up brand trolling would be an interesting future research avenue. Furthermore, comparing existing trolling scales, such as the Global Assessment of Internet Trolling (GAIT) scale (Buckels et al., 2014), and the prospective brand trolling scale would be worthwhile to see how the dimensions and items may differ between the constructs. Finally, examining whether brand trolling is unidimensional like other negative behaviors such as grieving (Ladanyi & Doyle-Portillo, 2017) would be an interesting area to explore. Overall, it would be valuable to determine the dimensionality and conceptualization of brand trolling, because it would make scholars better equipped to research the behavior in the future. Accordingly, we postulate the following research question:

RQ4: When developing a brand trolling behavior scale, what items and dimensions should be included in the scale to reflect what brand trolling accurately is?

Positive Brand Trolling

Trolling research has often discussed the positive aspects and outcomes of the behavior (e.g., Cruz et al., 2018). Knowing this, it would be of interest to determine whether brand trolling can also have positive characteristics and generate positive outcomes. For instance, previous research (Golf-Papez & Veer, 2017) has mentioned how trolling and negativity can lead to increased traffic to websites, increased interactions, closer communities (Coles & West, 2016b), increased sales through enhanced levels of product awareness (Berger et al., 2010), and increased product usage (Nepomuceno et al., 2023). These findings are not surprising given the fact that social media platforms tend to highlight negative and anger-inducing content to users in order to keep them engaged (Pelley, 2021). Similar to the adage of “any publicity is good publicity,” could the same sentiment apply to brand trolling in communities where any kind of engagement (in this case, trolling) would be seen as a positive for the community? It would be interesting to examine whether the aforementioned outcomes could apply in a brand community setting as well. Furthermore, the management of brand trolling could also inadvertently generate positive outcomes such as better customer attitudes and enhanced brand and community reputation if members perceive that the brand trolling was dealt with appropriately. Overall, it would be intriguing to examine whether brand trolling can generate positive outcomes in the same way that other contentious behaviors are able to. Accordingly, we postulate the following research question:

RQ5: What are the positive consequences that can result from brand trolling and brand trolling management?

Meta-Level Examination

The relationships and frameworks in this conceptual research were investigated on the micro-, meso- and macro-level where the effects on the individual, community and organization/brand were considered. However, can the investigation be expanded into the so-called meta-level where effects on the overall ecosystem are being examined? For instance, meta-level research could assess whether the prevalence of brand trolling could impact the overall number of communities, community members, and social media users in general. If brand trolling is highly common, it may cause users and members to leave online communities and social media platforms permanently. A common response to excessive social media use, which can generate negative consequences such as low self-esteem, trolling victimization, and health problems (Zahrai et al., 2022), has been social media detoxes, which is when individuals quit using social media to improve their well-being (El-Khoury et al., 2021). From a brand perspective, this exodus of users could negatively affect customer engagement in many areas, which could consequently have a negative impact on crucial organization technologies and processes such as big data analysis and customer relationship management (CRM). Accordingly, it would be interesting to determine whether brand trolling could also generate these types of far-reaching outcomes on the ecosystem level. From this, we postulate the following research question:

RQ6: What are the relevant predictors, outcomes and contextual elements of brand trolling on the meta-level (i.e., the ecosystem level)?

Trolling by Brands

Another potential fruitful area of research revolves around using trolling as a potential marketing tool. Recent research has suggested the possibility of organizations engaging in trolling themselves (e.g., Behl & Bhutani, 2022; Golf-Papez & Veer, 2022). This brand-initiated trolling could generate many potential research avenues. For instance, the research could focus on the potential goals that organizations may have in mind when deciding to engage in trolling, such as increasing engagement (i.e., likes, comments, and shares on social media) or increasing follower/subscriber numbers. Likewise, longitudinal research could examine how organizations conduct trolling at different points throughout the campaign and investigate the audience's response to the so-called *troll marketing*. Finally, interviews could be conducted with social media managers that delve deeper into the plans, tactics, strategies, and goals of *troll marketing*, which could provide valuable findings in this emerging area of trolling research. From this, we postulate the following research question:

RQ7: What are the strategies and goals of troll marketing?

Individual-Directed Trolling vs. Brand-Directed Trolling

Further investigating the key differences between general trolling directed at individuals and trolling directed at brands and brand-related actors would be another intriguing avenue of research. For instance, Demsar et al. (2021) investigated brand-directed trolling as a construct and found many similarities with general trolling directed at individuals. The authors found that both constructs can include malicious and negative practices, and both can be deceptive and provoke reactions. However, as we discovered in this research, it is clear that general trolling and brand trolling also differ in some crucial ways. As a further inquiry, it would be interesting to delineate the key differences between individual-directed trolling and

brand-directed trolling. It would be worthwhile to assess how these constructs differ with regards to their predictors, consequences, contextual elements, or any other relevant facets. Accordingly, this leads to the following research question:

RQ8: When directed towards a brand or an organization instead of an individual, how does trolling differ in its relationships, contextual elements, frequency, severity and manifestation?

Furthermore, examining the differences between the perceptions of individual-directed trolling and brand-directed trolling would be an appealing avenue of research. Rotman et al. (2017) discovered how individuals are willing to engage in unethical behavior to punish harmful brands. This finding reveals how individuals are sometimes willing to engage in behaviors such as trolling because they feel the target of the behavior deserves it. If an organization or a brand has engaged in unethical behaviors, then it is possible that the public may view the trolling individual or group in a more understanding light. Additionally, it is possible that individuals may feel less compassion for impersonal brands compared to humans. From this, it would be interesting to examine how the online public would view brand-directed trolling compared to individual-directed trolling. Consequently, we postulate the following research question:

RQ9: How would the public's perceptions differ when viewing brand-directed trolling (where the targets include organizations and brands) compared to individual-directed trolling?

Management of Brand Trolling

The management of brand trolling is another key area of research that could be further investigated. Compared to general trolling, it was found that community members, targets, employees and the organizations themselves are better equipped and more inclined to manage brand trolling taking place in a brand community. However, managing the behavior will also impact the actors and community, which means it is vital to strike an appropriate balance between a fully zero-tolerance and *laissez-faire* approach. This raises some interesting research questions about the outcomes of managing brand trolling on the community- and brand-level. For instance, can troll management tactics and strategies generate the same outcomes as the trolling behavior itself? More specifically, is it possible for trolling management strategies to have an effect on attitudes, joining behaviors, engagement, satisfaction and other community- and brand-related outcomes? Drawing on RP10 and RP11, the following research questions emerge:

RQ10: What is the relationship between trolling management strategies and community-related factors such as community members' (a) joining behaviors (b) level of engagement, (c) quality of engagement, (d) brand community satisfaction levels, (e) visiting and (f) revisiting intentions, (g) community trust levels, and (h) positive WOM intentions towards the brand community?

RQ11: What is the relationship between trolling management strategies and brand-related factors such as (a) customer brand attitudes, (b) brand reputation, (c) positive WOM intentions towards the brand, (d) brand image, (e) brand trust, (f) brand satisfaction, and (g) brand purchase intentions?

Conclusions

This investigation reveals how trolling is a common online behavior that has become much more complex than originally thought. Trolling can be directed at individuals, online groups and communities, products, brands and even organizations. The multi-leveled, relational and contextual nature of brand trolling leads to a plethora of fundamental research propositions and questions that focus on a wide range of unique actors, contexts and relationships. This research clearly demonstrates how there is a lot to be examined and discovered with regards to brand trolling.

Overall, *Essay I* provides a more comprehensive understanding of trolling in a marketing context. More specifically, this is accomplished by introducing novel actors beyond the troll-target dyad, such as other community members and employees. Furthermore, the dynamic relationships and contextual elements of brand trolling aid scholars in seeing its more holistic nature where its influence is far-reaching to several actors and institutions. Finally, examining the proposed predictors and outcomes of brand trolling as well as its contextual elements, can potentially help future decision-makers and managers better understand how the behavior emerges and what can be done to minimize its negative effects.

This conceptual research aids in the development of a solid foundation for brand trolling. If the research propositions and questions are properly investigated and answered, then brand- and organization-related stakeholders will better understand brand trolling and, therefore, will be better equipped to manage the behavior. Consequently, this leads to brand communities being better able to serve their original purpose, which is about encouraging fruitful participation, interactions and content creation among dedicated brand enthusiasts.

To conclude, it is the hope of the author that the findings in this essay prompt further inquiry into the brand trolling construct so that it can be more properly managed by relevant stakeholders in the future. Further investigation into the research propositions and questions of *Essay I* as well as any other relevant research propositions suggested in previous trolling literature will ideally lead to a better overall comprehension of trolling and brand trolling behavior and victimization. However, before these empirical queries can be answered, it is necessary to understand what trolling exactly is and what it entails as a construct from both the troll and target perspectives. Fittingly, to accomplish these goals, we will shift to *Essay II*, which focuses on developing a comprehensive conceptualization of trolling behavior and trolling victimization through following the key steps and processes of scale development.

Essay II: Scale Development for Trolling Behavior and Victimization

Abstract

Trolling is a behavior that has become increasingly common in many online contexts. It has become so prevalent that online users are usually able to recognize trolling when it happens, however, when asked, these same users are often hard-pressed to define what trolling exactly is. Overall, there is some confusion regarding the exact conceptualization of trolling behavior from online users, practitioners and scholars alike. Accordingly, this research addresses this issue by gaining a more comprehensive understanding of trolling behavior and its impact on a wide range of actors. This was accomplished by developing valid and reliable scales for both trolling behavior and trolling victimization. The results of exploratory and confirmatory factor analyses, reliability tests and multiple validity tests have confirmed that both trolling behavior and victimization are reliable, valid, and multidimensional constructs that contain factors of distraction, deception, and provocation.

Keywords: Online Trolling, Trolling Victimization, Scale Development, Online Communities, Measurement, Psychometrics

Introduction

Online technologies and social media have consistently grown in prominence and influence in recent years due to their ability to promote highly social online environments (Obar & Wildman, 2015) and communities where individuals can instantly communicate and interact with other individuals, groups, businesses, and brands. By adopting online communities, brands benefit from the communities fostering long-term customer-brand relationships and generating positive outcomes such as enhanced financial performance and customer attitudes (Baldus et al., 2015). Furthermore, online communities benefit the members by providing in-depth brand-related information and knowledge, generating unique brand experiences (Baldus et al., 2015), and creating a central hub where ideas, solutions, and discussions among devoted brand enthusiasts can occur.

However, the instantaneous nature of online participation can act as a double-edged sword for brands and users, creating opportunities for both positive and negative engagement. Negative forms of online engagement, such as trolling, have become increasingly prevalent in online contexts (Golf-Papez & Veer, 2017; YouGov, 2014). Trolling is an undirected and unprovoked behavior that consists of deceptive, disruptive, and deliberate actions that occur in online settings that intend to provoke negative reactions from individuals (Dineva & Breitsohl, 2021; Sanfilippo et al., 2018).

Trolling is a problematic behavior that has been linked to several negative outcomes such as increased hostility (Rösner et al., 2016), feelings of anger, aversion, and hatred (Anderson et al., 2014; Gervais, 2015), reduced satisfaction with online platforms (Gervais, 2015), and an increase in the likelihood of trolling occurring (Cheng et al., 2017). Trolling can also have an impact on a wider scale where it negatively affects the brands and online communities where it takes place. Trolling has been linked to outcomes such as users leaving sites (McAloon, 2015), harmed brand images and reputation, lower levels of employee well-being, and financial losses to companies (Demsar et al., 2021; Golf-Papez & Veer, 2017).

While trolling literature has gained traction recently, the construct and its impact has been understudied in several contexts, including marketing (Golf-Papez & Veer, 2017). Furthermore, there has been little research on trolling victimization; most research about online victimization has focused on the victims of cyberbullying (e.g., Patchin & Hinduja, 2010). The research on trolling and its victimization may be scarce because trolling lacks a proper conceptual foundation and is often perceived as an umbrella term that is used interchangeably with similar behaviors (Jussinoja, 2018; Marwick & Lewis, 2017; Sun & Fichman, 2019). Furthermore, trolling is a highly contextual and changing behavior that not only differs across communities, platforms, and cultures (Sanfilippo et al., 2018; Sun & Fichman, 2019; Synnott et al., 2017), but it also differs in its manifestations, meanings and effects (Dyrel, 2016; Fichman & Peters, 2019). These factors explain some of the conceptual confusion and misuse surrounding the trolling construct.

These aforementioned findings provide some initial evidence of how trolling is yet to be fully understood by scholars and practitioners, which can pose challenges for both actors. Accordingly, valid and reliable scales for trolling behavior and victimization must be developed before empirically researching how trolling can impact online platforms and communities. This is a crucial and necessary step because creating appropriate measurement scales contributes to a more comprehensive understanding of trolling and its potential impact on relevant actors and groups. If trolling and its impact can be better comprehended and studied, then marketing practitioners can better identify and manage the behavior in their communities. As a result, the communities will be more harmonious and unified in their interactions and discussions. Finally, generating a suitable conceptualization of trolling means

that scholars would be better equipped to appropriately research trolling behavior and its victimization in the future.

The remainder of this essay will be structured as follows: Firstly, a literature review will be conducted to determine the current findings on the conceptualizations of trolling behavior and victimization. Next, the scale development process will be conducted, and the results from the pre-test, expert review, and studies will also be presented and discussed. Following this, a discussion of the relevant theoretical and practical implications will occur. Finally, limitations and future research avenues will be examined to conclude the essay.

Literature Review

Trolling Behavior & Victimization

Current Research

Trolling is a phenomenon that has become increasingly prevalent in recent literature, especially in the psychology and information technology context (e.g., Howard et al., 2019; Maltby et al., 2016; Paavola et al., 2016; Thacker & Griffiths, 2012). Trolling literature has examined topics such as predictors (Howard et al., 2019), motivations (Cook et al., 2017), response strategies (Cook et al., 2017), dimensions (Sanfilippo et al., 2018), and consequences (Buckels et al., 2018). However, as previously mentioned, trolling has received little attention in marketing literature, barring a few exceptions (e.g., Demsar et al., 2021; Golf-Papez & Veer, 2017, 2022).

While there has been increasing research on trolling behavior, there has been relatively little attention on trolling victimization specifically. Hong and Cheng (2018) examined different forms of trolling victimization and found that a sense of inferiority and depression were significant predictors of victimization. Furthermore, trolling victimization was found to be a significant predictor of trolling behavior (Wilson & Seigfried-Spellar, 2022), which falls in line with the findings of previous research where victims of cyberbullying were more likely to engage in cyberbullying later on (Patchin & Hinduja, 2011). However, most victimization research tends to focus on victims of cyberbullying where topics such as consequences (Patchin & Hinduja, 2010), predictors (Guo, 2016), moderators (Guo, 2016) and scale development (Çetin et al., 2011; Stewart et al., 2014; Thomas et al., 2019) have been investigated.

The scant research on both trolling behavior and victimization may be explained by the fact that trolling lacks conceptual clarity (March & Marrington, 2019). Trolling is often a highly deceptive and context-dependent behavior, which creates confusion for practitioners, victims, and scholars alike, because it is challenging to determine when the behavior is actually occurring. This confusion is demonstrated in trolling research where scholars often create their own interpretations of trolling (e.g., Buckels et al., 2014; Hardaker, 2010; Jane, 2015; March & Marrington, 2019). This is problematic, because it illustrates how trolling lacks a conceptual foundation where there is no agreement on its exact definition and conceptualization, which leads to disparate findings and further obfuscation regarding the construct.

Current Scales

Despite the issues regarding its conceptualization, there are some existing scales on trolling behavior. For instance, the 4-item Global Assessment of Internet Trolling (GAIT) scale by Buckels et al. (2014), and the 12-item iTroll questionnaire by Buckels et al. (2018) measure trolling behavior, enjoyment, and identification. Also, Maltby et al. (2016) created a scale

called the “Conceptions of Online Trolls” to assess if certain group factors were associated with resilience to trolling. Hong and Cheng (2018) generated scales for online trolling behavior and victimization specifically for university students and they found four different types of trolling and victimization.

There are a few issues that are present in the current trolling scales. Firstly, the scales are often narrow in scope where they use a particular sample (e.g., students) and they are specifically developed to answer the questions set forth in the research; they are not necessarily developed for broader use by other scholars. Secondly, the scales measuring trolling are often unidimensional, which diverges from the existing definitions of trolling that often include underlying factors such as deception, disruption, provocation, aggression, and maliciousness (Buckels et al., 2014; Craker & March, 2016; Hardaker, 2010; Jane, 2015; March & Marrington, 2019). This shows how the existing conceptualizations of trolling behavior and victimization are somewhat disjointed and not yet fully comprehended by scholars. If trolling behavior is fully understood, it would be easier for relevant organization-related actors to identify and manage the behavior in their online platforms and communities. If trolling victimization is not fully understood, then it would be more difficult for those same actors to resolve the issues and maintain their relationships with the victims of trolling. Finally, the lack of comprehensive knowledge on trolling and victimization would make it more challenging for scholars to research both constructs appropriately. This demonstrates the need for comprehensive scales that better encapsulate the trolling behavior and victimization constructs.

Simultaneous Measuring of Trolling and Victimization

It is important to note the challenging nature of simultaneously developing scales for two separate constructs. However, scales for negative behaviors and victimization have been simultaneously developed before with success. For instance, Çetin et al. (2011) simultaneously examined cyberbullying behaviors and victimization and found a three-factor solution for each construct. The factors were cyber forgery, cyber verbal bullying, and hiding identity, and they were all reliable and valid (Çetin et al., 2011). Furthermore, Thomas et al. (2019) simultaneously developed bullying and cyberbullying victimization and perpetration scales. The scales were reliable and showed evidence of concurrent and convergent validity (Thomas et al., 2019). These studies demonstrate that it is possible to develop two constructs simultaneously.

Trolling vs. Cyberbullying

One emerging theme from the literature review is how trolling behavior and victimization have some clear parallels with cyberbullying behavior and victimization. This could lead to some confusion, especially when trolling is often viewed as an umbrella term for negative online behaviors (Jussinoja, 2018; Marwick & Lewis, 2017; Sun & Fichman, 2019). Therefore, it is crucial to delineate how trolling is a unique and distinct construct that is separate from similar constructs such as cyberbullying.

The literature on cyberbullying suggests that cyberbullying is an intentional, direct, and targeted behavior (Steffgen et al., 2011). Furthermore, cyberbullies and their victims often know one another in real life (Dooley et al., 2010). In comparison, trolling has been viewed as a more random behavior that is typically not directed at any particular individual (Craker & March, 2016; de Seta, 2013; Golf-Papez & Veer, 2017). Trolling can also be humorous and playful at times, whereas cyberbullying is mainly intended to be harmful and discomfiting (Tokunaga, 2010). Finally, trolling can include distinct behaviors such as deception and disruption, whereas cyberbullying does not include those kinds of behaviors (Craker & March,

2016). To provide evidence of trolling being distinguishable from cyberbullying, Study 3 will include trolling and cyberbullying perpetration and victimization scales to assess the discriminant validity of the constructs.

Dimensions of Trolling

An extensive content analysis was conducted to determine the potential underlying factors of trolling behavior and victimization. 212 journal articles and research papers that included definitions and conceptualizations of trolling were investigated to determine the initial measurement items and the core underlying factors of the constructs. The three most common factors found in the analysis were provocation (86 mentions), disruption (79 mentions), and deception (38 mentions). These were chosen to be the main factors of trolling behavior and victimization, and they will be discussed in greater detail in the following paragraphs. It is important to note that these factors apply to both trolling behavior and victimization, because there will always be a perpetrator and victim when a trolling incident occurs.

Provocation is the first factor of trolling that will be discussed. Provocation refers to deliberate behaviors intended to be inflammatory and make others annoyed or angry (Sanfilippo et al., 2018). More specifically, provocation is about eliciting negative emotional reactions and responses from other individuals through interacting with them in an aggressive and/or contentious manner (Sanfilippo et al., 2018). Essentially, the purpose of this factor is to instigate other individuals into retaliation by being controversial and/or overly argumentative in one's actions.

Next, disruption refers to aggravating, irrelevant, and repetitive behavior that has no practical purpose (Hardaker, 2010). More specifically, this attention-seeking behavior is intended to distract from the actual purpose of the online community by generating confusion and/or unfocused responses (Cruz et al., 2018; Hardaker, 2010; Leaver, 2013). This behavior is not necessarily about attacking one specific individual, rather it is about having a broader impact and diverting the attention of as many individuals as possible.

Finally, deception refers to deliberately sending messages to an individual "with the intent to foster a false belief or conclusion" (Caspi & Gorsky, 2006). In a trolling context, deception can manifest in a few ways. Firstly, there is identity deception where the trolls misrepresent themselves and attempt to be viewed as a genuine user (Cruz et al., 2018). For instance, this can be accomplished by creating fake profiles (Buelga et al., 2020). Secondly, deception can also be about misdirecting others where the troll's participation does not reflect how they truly feel, rather the purpose of their participation is to upset and gain strong reactions from others (Chen, 2018; Cook et al., 2019). Finally, deception can also emerge as fake reviews that intend to mislead customers in purchasing decisions (Zhang et al., 2016). The deception factor demonstrates the more relational nature of trolling (Ellcessor, 2016), because trolls must understand the nuances and mannerisms of their audience to be perceived as legitimate.

Next, the methods used to manage common method bias and generate and purify the items that will measure the constructs and generate the scales for trolling behavior and victimization will be discussed.

Common Method Bias

It is important to note that some methodological challenges emerge when measuring negative behaviors such as trolling. More specifically, when developing cross-sectional studies that rely on self-reported data, it is possible for common method bias to emerge (Lindell & Whitney, 2001; Podsakoff et al., 2012). Essentially, if the same respondent in a study is

simultaneously providing answers for predictor and outcome constructs, then this may lead to inflated correlations between the constructs (Podsakoff et al., 2012).

To address the potential issue of common method bias, some procedural and statistical solutions recommended by previous research (e.g., Carpenter, 2017; Olson, 2010; Podsakoff et al., 2003, 2012) were introduced. The procedural methods will first be discussed. Firstly, the measurement items were tested in pre-tests and expert reviews to minimize any confusion or ambiguity that respondents may have had with regards to survey item and construct comprehension. Second, to reduce evaluation apprehension, respondents were provided with a guarantee of anonymity and confidentiality with regards to their responses. Furthermore, when providing answers for certain constructs, some pages clearly displayed the following message “There are no right or wrong answers, we are interested in your honest answers. Remember, your confidentiality is protected as all responses are anonymous and cannot be tied to your identity”. Finally, to control for potential item-order effects, the items for the outcome constructs were presented before the predictor constructs.

Some statistical remedies were also taken to minimize the likelihood of common method bias. Firstly, a confirmatory factor analysis was conducted to compare a one-factor model against the hypothesized multi-factor model. For both trolling behavior and trolling victimization, the results from the Study 2 and 3 samples indicated that the fit statistics for the one-factor models were significantly worse compared to the multi-factor models (see Tables 13 and 14 for Study 2 results), which provides initial evidence of a lack of common method bias. Next, using the Study 3 sample, the correlation coefficients between the key trolling behavior constructs (distraction, deception, provocation, sadism, depression, psychopathy, agreeableness) ranged from -0.552 to 0.759. For trolling victimization, the correlation coefficients between the key constructs (distraction, deception, provocation, hostility, depression, a willingness to divulge personal information, introversion) ranged from 0.030 to 0.671. This provides further evidence of a lack of common method bias, because all values were below the recommended threshold of 0.90 (Pavlou et al., 2007). Overall, this aggregate evidence demonstrates that the probability of common method bias affecting the results of the studies is quite low.

Item Generation and Pre-Test

This research followed the procedures and recommendations from established scale development paradigms (e.g., Churchill, 1979; Gerbing & Anderson, 1988) and guidelines (e.g., Boateng et al., 2018; Carpenter, 2017; Worthington & Whittaker, 2006).

Based on the aforementioned content analysis of 212 research papers, a preliminary pool of 44 items was generated for both trolling behavior and victimization. The three most commonly mentioned factors of provocation, disruption and deception were included in the first iteration of the survey with provocation initially containing 18 items, distraction containing 13 items, and deception containing 13 items. The items were measured on a 5-point Likert scale assessing frequency ranging from “Never” (1) to “Always” (5). The items measuring trolling behavior began with the statement “*I have deliberately...*” followed by the action itself (e.g., caused distractions online). The items measuring trolling victimization began with the statement “*I have experienced individuals deliberately...*” followed by the action (e.g., causing distractions online). Also, some background questions about demographics were asked. The survey was pre-tested to determine if any items should be changed, removed, or added. Respondents were also debriefed at the end of the survey to provide feedback and comments on the quality of the survey. In total, 57 responses were collected for the victimization pre-test, and 53 responses were collected for the trolling behavior pre-test, which is in line with the expected sample size of 5 to 100 for pre-tests (Carpenter, 2017). Some minor

changes in language were made to the items and questions to improve readability and provide further clarity for respondents. However, no items were removed or added based on the pre-test results.

Expert Review

An expert review was conducted to enhance the robustness of the survey by providing evidence of face validity. Previous research has stated that five to seven experts is an appropriate number of respondents for an expert review (Boateng et al., 2018). In total, eight experts provided feedback. The experts consisted of scholars who have experience with researching trolling and other similar negative online behaviors. Their task was to assess the representativeness of each item on a 3-point scale, provide feedback on the relevance and quality of the survey items, and determine whether any items should be changed, removed, or added. An item-rating task was used to determine the representativeness of the items in the survey. Consistent with previous research (e.g., Busser & Shulga, 2018), if the item received a score of 75% or higher (i.e., an average score of 2.25 out of 3), then the item would be retained. In total, six items were removed based on the item-rating task. Furthermore, the language and wording in some items and questions were changed to improve readability and provide more clarity for the respondents. Finally, based on expert feedback, four items were added to the survey, which led to a total of 42 items (with 13 distraction items, 9 deception items and 20 provocation items) being retained for both trolling behavior and victimization. The final individual measurement items for each dimension can be found in Appendices C to H in the Appendices section.

Study 1

Sample

The purpose of Study 1 was to further refine and purify the scale, which was accomplished with an exploratory factor analysis (EFA). Firstly, a short qualifier survey was created to ensure that respondents have actually engaged in trolling behavior or been victims of trolling behavior. In total, 495 responses were collected for the trolling survey, and 427 responses were collected for the victimization survey. The surveys were developed on the Qualtrics platform. The respondents were from either the United States or Canada, and they were recruited from Amazon Mechanical Turk. Respondents were compensated \$1.15 for completing both the qualifier and main survey.

After cleaning the data for careless responding, 390 responses were retained for the trolling survey, and 389 responses were retained for the victimization survey. Responses were flagged and considered for removal based on established methods such as fast response detection (Curran, 2016; Greszki et al., 2015), low intra-individual response variability (IRV) detection (Dunn et al., 2018), long-string analysis (Meade & Craig, 2012), outlier analysis with Mahalanobis distances (Curran, 2016), and missed attention check detection (Curran, 2016). Responses were removed when three or more of these issues were flagged, straightlining was detected, or the average response time to items was 2 seconds per item or faster (Huang et al., 2012). The sample size for both surveys met the minimum requirement of 200-300 cases and the 5:1 respondent-to-item ratio (Boateng et al., 2018; Carpenter, 2017, Worthington & Whittaker, 2006). The demographic characteristics of both sets of respondents are summarized in Tables 4 and 5.

Table 4: Demographic Characteristics of Study 1 Respondents (Trolling Behavior)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	5	1.3%
United States	385	98.7%
Biological Gender		
Male	218	55.9%
Female	171	43.8%
Non-binary/Prefer to self-describe	1	0.3%
Age		
18-24 years old	17	4.4%
25-34	173	44.4%
35-44	116	29.7%
45-54	44	11.3%
55-64	32	8.2%
+65	8	2.1%

Table 5: Demographic Characteristics of Study 1 Respondents (Trolling Victimization)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	4	1.0%
United States	385	99.0%
Biological Gender		
Male	175	45.0%
Female	211	54.2%
Non-binary/Prefer to self-describe	3	0.8%
Age		
18-24 years old	21	5.4%
25-34	149	38.3%
35-44	120	30.8%
45-54	53	13.6%
55-64	37	9.5%
+65	9	2.3%

EFA Findings

Factorability of Data

After gathering the responses, the factorability of the data was first assessed (Dziuban & Shirkey, 1974). The Kaiser-Meyer-Olkin (KMO) value exceeded 0.60 for trolling behavior (0.959) and victimization (0.945). Bartlett's chi-square was also significant for trolling behavior (<0.001) and victimization (<0.001). Finally, the correlation matrices for both constructs included several values that were higher than 0.30. These findings suggest that factor analysis was appropriate for the data.

Factor Solution

Next, an EFA was run with principal axis factoring using promax rotation. Items were removed due to low factor loadings (<0.32) (Hair et al., 2019), low communalities (<0.50), and high cross-loadings where there was "less than a 0.15 difference from an item's highest factor loading" (Worthington & Whittaker, 2006, p. 823). The analysis generated a three-factor solution for both trolling behavior and victimization (where eigenvalues >1). For trolling behavior, 17 items were retained and the solution explained 65.89% of total variance. For trolling victimization, 18 items were retained and the solution explained 63.27% of total variance. Furthermore, the three-factor solution is justified because prior literature has stated that stopping criteria for factor retention can include "a predetermined number of factors based on research objectives and/or prior research" (Hair et al., 2010, p. 111), and having enough factors to meet a certain percentage of variance explained, usually 60% or higher (Hair et al., 2010). Next, the factor loadings were all above 0.40 for both trolling behavior (see Table 6) and trolling victimization (see Table 7). The potential issue of multicollinearity was assessed by examining the determinant of the correlation matrix (Field, 2018; Tabachnick & Fidell, 2013). The determinant was 0.000026 for trolling behavior and 0.0000351 for victimization, which was above the threshold value of 0.00001, which means multicollinearity was not an issue for either construct.

Table 6: Factor Loadings for 17-item Trolling Behavior Solution (EFA)

Items	Factor 1	Factor 2	Factor 3
Caused distractions	0.641		
Caused inconveniences for individuals	0.629		
Caused disruptions	0.606		
Caused confusion	0.592		
Caused disturbances	0.713		
Caused dysfunction	0.833		
Lured individuals into pointless arguments	0.661		
Misrepresented myself			0.863
Provided false information			0.634
Lied to individuals			0.658
Created false profiles			0.720

Items	Factor 1	Factor 2	Factor 3
Bothered individuals		0.652	
Made individuals react negatively		0.742	
Provoked individuals		0.620	
Made individuals respond negatively to me		0.644	
Been overly argumentative with individuals		0.748	
Been controversial in communication		0.683	

Table 7: Factor Loadings for 18-item Trolling Victimization Solution (EFA)

Items	Factor 1	Factor 2	Factor 3
Causing distractions		0.808	
Causing inconveniences for others/me		0.711	
Causing disruptions		0.742	
Causing confusion		0.521	
Causing disturbances		0.657	
Causing dysfunction		0.713	
Creating disorder		0.753	
Misrepresenting themselves			0.413
Providing false information			0.505
Creating false profiles			0.853
Creating false reviews			0.810
Being inflammatory to others	0.694		
Mocking others/me	0.880		
Making others/me react negatively	0.682		
Provoking others/me	0.768		
Being controversial in communication	0.653		
Causing arguments	0.714		
Creating conflict	0.758		

Reliability

Finally, reliability analyses were conducted for all three factors from both constructs. The Cronbach's alpha values ranged from 0.827 to 0.906 (see Table 8), meaning they were all above the threshold value of 0.70 and, therefore, reliability was confirmed (Hair et al., 2019).

Table 8: Reliability Statistics for Trolling and Victimization Factors

Factor	Cronbach's Alpha
Trolling: Distraction	0.903
Trolling: Deception	0.860
Trolling: Provocation	0.888
Victimization: Distraction	0.885
Victimization: Deception	0.827
Victimization: Provocation	0.906

Study 2

The purpose of Study 2 was to validate and confirm the scale, which was accomplished with a confirmatory factor analysis (CFA). In addition, known-group validity and predictive validity were also assessed in this study.

Sample

Similar to Study 1, a short qualifier survey was created to ensure that respondents have actually engaged in trolling behavior or been victims of trolling behavior. In total, 367 responses were collected for the trolling survey, and 441 responses were collected for the victimization survey. The surveys were developed on the Qualtrics platform. The respondents were from either the United States or Canada, and they were recruited from Amazon Mechanical Turk. Respondents were compensated \$0.95 for completing both the qualifier and main survey (Note: The lower compensation was due to the shorter length of the survey).

After cleaning the data for careless responding in the same manner as in Study 1, 307 responses were retained for the trolling survey, and 348 responses were retained for the victimization survey. The demographic characteristics of both sets of respondents are summarized in Tables 9 and 10.

Table 9: Demographic Characteristics of Study 2 Respondents (Trolling Behavior)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	3	1.0%
United States	304	99.0%
Biological Gender		
Male	169	55.0%
Female	137	44.6%
Non-binary/Prefer to self-describe	1	0.3%
Age		
18-24 years old	27	8.8%
25-34	161	52.4%
35-44	70	22.8%
45-54	28	9.1%
55-64	20	6.5%
+65	1	0.3%

Table 10: Demographic Characteristics of Study 2 Respondents (Trolling Victimization)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	5	1.4%
United States	343	98.6%
Biological Gender		
Male	189	54.3%
Female	158	45.4%
Non-binary/Prefer to self-describe	1	0.3%
Age		
18-24 years old	18	5.2%
25-34	185	53.2%
35-44	74	21.3%
45-54	46	13.2%
55-64	22	6.3%
+65	3	0.9%

CFA Findings

A confirmatory factor analysis (CFA) was conducted to verify the factor structure of the observed variables. The preferred approach for CFA was structural equation modelling (SEM), which was completed using the AMOS 29 software. CFAs were run on the 17-item

and 18-item solutions for trolling behavior and trolling victimization respectively. Initially, the three-factor first-order models had good fit statistics for both constructs. Based on the recommendations of prior research (Ahmad et al., 2016; Bagozzi & Yi, 1988), the three-factor models were further refined by removing items one at a time due to high modification indices. As a result, two items were removed from both scales due to high modification indices. For trolling behavior, the “*Lured individuals into pointless arguments*” and “*Bothered individuals*” items were removed. For trolling victimization, the “*Misrepresenting themselves*” and “*Being inflammatory to others/me*” items were removed.

Fit Statistics & Model Statistics

Following this, the fit statistics were evaluated once more. It is important to note that the fit statistics contain indices and values suggested by previous research (Hair et al., 2010) including the χ^2/df value, an absolute fit index (e.g., GFI, RMSEA, or SRMR), an incremental fit index (e.g., CFI or TLI), a goodness-of-fit index (e.g., GFI, CFI, or TLI) and a badness-of-fit index (e.g., RMSEA or SRMR). For trolling behavior, the following fit statistics were reached: ($\chi^2/df = 2.029$, $p = 0.000$, RMSEA = 0.058, CFI = 0.970, TLI = 0.964, SRMR = 0.0300). For trolling victimization, the following fit statistics were achieved: ($\chi^2/df = 1.949$, $p = 0.000$, RMSEA = 0.052, CFI = 0.967 TLI = 0.961, SRMR = 0.0325). The fit statistics for both models met the threshold values determined by previous research (e.g., Hair et al., 2010), which provided evidence of good-fitting models. The p -values were significant at 0.000 for both models, which was an issue, however, this is to be expected with models where the number of respondents is greater than 250 and the number of variables is between 12 and 30 (Hair et al., 2010). Also, the factor loadings ranged from 0.717 to 0.839 for trolling behavior (see Table 11), and from 0.677 to 0.806 for trolling victimization (see Table 12), which is above the minimum threshold of 0.50 (Hair et al., 2010). Finally, an assessment of the normality of the data was completed. The skewness and kurtosis values were between -2 and +2 for both models, which provided evidence of normality in the data (Lomax & Hahs-Vaughn, 2012).

Table 11: Factor Loadings for 15-item Trolling Behavior Solution (CFA)

Items	Distraction	Deception	Provocation
Caused distractions	0.786		
Caused inconveniences for individuals	0.812		
Caused disruptions	0.839		
Caused confusion	0.739		
Caused disturbances	0.815		
Caused dysfunction	0.832		
Misrepresented myself		0.787	
Provided false information		0.778	
Lied to individuals		0.818	
Created false profiles		0.763	
Made individuals react negatively			0.783
Provoked individuals			0.791

Items	Distraction	Deception	Provocation
Made individuals respond negatively to me			0.717
Been overly argumentative with individuals			0.743
Been controversial in communication			0.727

Table 12: Factor Loadings for 16-item Trolling Victimization Solution (CFA)

Items	Distraction Victim	Deception Victim	Provocation Victim
Causing distractions	0.683		
Causing inconveniences for others/me	0.703		
Causing disruptions	0.757		
Causing confusion	0.677		
Causing disturbances	0.739		
Causing dysfunction	0.757		
Creating disorder	0.777		
Providing false information		0.702	
Creating false profiles		0.806	
Creating false reviews		0.742	
Mocking others/me			0.761
Making others/me react negatively			0.750
Provoking others/me			0.711
Being controversial in communication			0.750
Causing arguments			0.728
Creating conflict			0.763

Model Comparisons

Next, a comparison of models was done to determine which model had the best fit. The models to be compared were the original three-factor first-order model, a constrained model where the covariance between the factors was equal to one, a one-factor model where all items for trolling behavior and trolling victimization loaded onto one factor, and a second-order model where the factors of provocation, distraction and deception loaded onto a common factor of trolling (and victimization). Table 13 provides the fit statistics between the models for trolling behavior and Table 14 provides the fit statistics between the models for trolling victimization.

Table 13: Models and Fit Statistics for Trolling Behavior

Model	χ^2/df	RMSEA	CFI	TLI	SRMR
One-factor model	4.373	0.105	0.899	0.883	0.0515
Three-factor model	2.029	0.058	0.970	0.964	0.0300
Second-order model	2.062	0.059	0.969	0.963	0.0345
Constrained model	2.224	0.063	0.963	0.957	0.0731

Table 14: Models and Fit Statistics for Trolling Victimization

Model	χ^2/df	RMSEA	CFI	TLI	SRMR
One-factor model	4.063	0.094	0.892	0.875	0.0543
Three-factor model	1.949	0.052	0.967	0.961	0.0325
Second-order model	1.914	0.052	0.968	0.963	0.0329
Constrained model	2.278	0.061	0.955	0.948	0.1154

The model comparison provides evidence of how both trolling behavior and trolling victimization are clearly multidimensional constructs. The one-factor model had significantly worse fit statistics compared to the three-factor and second-order models. The constrained model had better fit statistics compared to the one-factor model, however, it had somewhat worse fit statistics compared to the three-factor and second-order models. Also, the results showed a significant difference between the unconstrained three-factor model and the constrained model for both trolling behavior ($\Delta\chi^2 = 23.683$, $\Delta df = 3$, and $p = 0.000$), and trolling victimization ($\Delta\chi^2 = 40.033$, $\Delta df = 3$, and $p = 0.000$), which provided initial evidence of discriminant validity. Also, the second-order models for trolling behavior and victimization (see Figures 2 and 3) had very similar fit statistics in comparison to the three-factor models, which shows how both models are good-fitting. Furthermore, the tenability of the second-order models (for trolling behavior and trolling victimization) was confirmed after running chi-square difference tests (Kline, 2016).

Figure 2: Trolling Behavior Second-Order Model

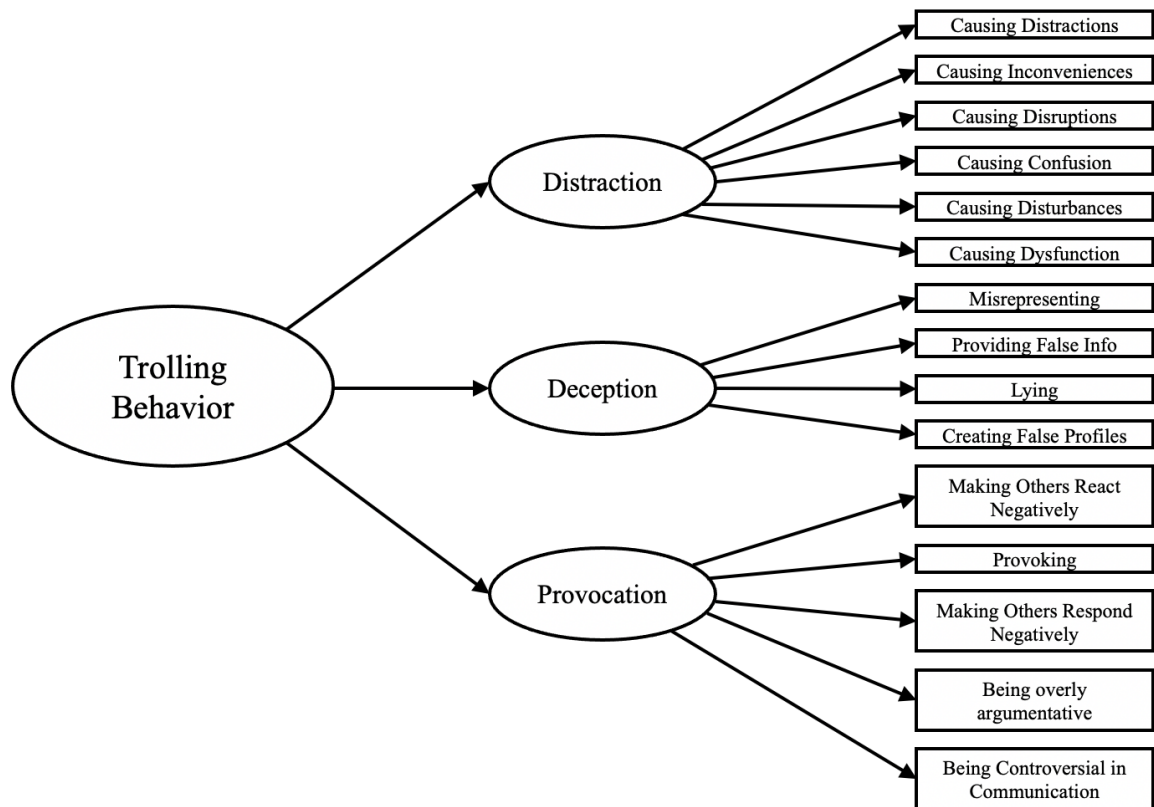
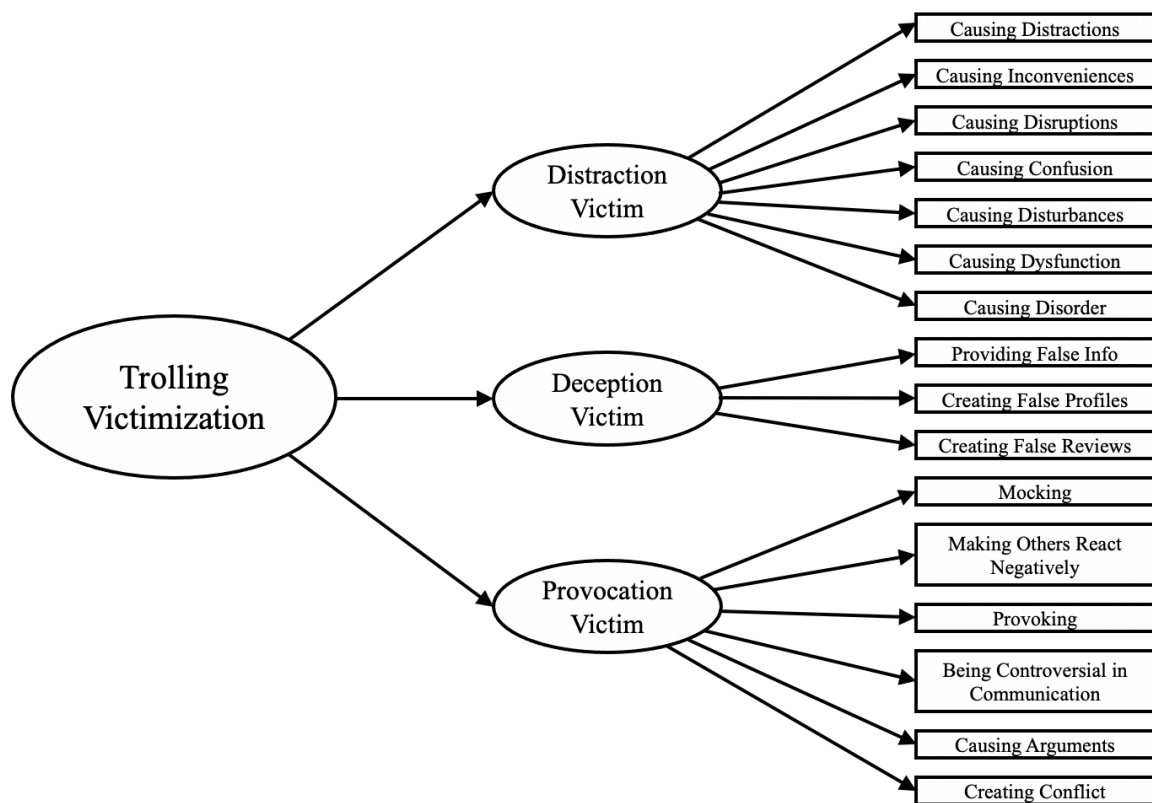


Figure 3: Trolling Victimization Second-Order Model



Reliability, Convergent Validity, and Discriminant Validity

Next, the reliability, convergent validity, and discriminant validity of both constructs was tested. Reliability was assessed with the composite reliability (CR) scores of each factor, whereas convergent validity was tested by examining the average variance extracted (AVE) for each factor. The CR values for all factors were above the threshold value of 0.70 (Hair et al., 2010) (see Tables 15 and 16). Furthermore, the AVE values were above the threshold value of 0.50 for all factors (Hair et al., 2010) (see Tables 15 and 16). These findings confirm the reliability and convergent validity of the trolling behavior and victimization factors.

Table 15: CR and AVE Values of Trolling Behavior Factors

Factor	Composite Reliability	Average Variance Extracted
Deception	0.867	0.619
Distraction	0.917	0.647
Provocation	0.867	0.567

Table 16: CR and AVE Values of Trolling Victimization Factors

Factor	Composite Reliability	Average Variance Extracted
Deception Victim	0.795	0.564
Distraction Victim	0.888	0.531
Provocation Victim	0.881	0.554

Next, discriminant validity was assessed. Typically, this has been accomplished by following the Fornell and Larcker (1981) criteria, which is done by determining that the shared variance between constructs is inferior to the AVEs of the individual measures. However, recent literature (Hair et al., 2017; Henseler et al., 2015) has suggested that this criterion may not reliably detect discriminant validity issues. Accordingly, the more recent and preferred method of assessing discriminant validity is by examining the heterotrait-monotrait (HTMT) ratio of the correlations, which is essentially the ratio of the between-trait correlations to the within-trait correlations (Henseler et al., 2015), where the threshold value of 0.90 must not be exceeded. The HTMT ratios for all factors of trolling behavior and victimization can be seen in Tables 17 and 18 respectively. None of the ratios for either construct exceeded the threshold value of 0.90 (Henseler et al., 2015), therefore, discriminant validity was confirmed.

Table 17: HTMT Ratios for Trolling Behavior Factors

	Provocation	Deception	Distraction
Provocation	-	0.850	0.861
Deception	0.850	-	0.802
Distraction	0.861	0.802	-

Table 18: HTMT Ratios for Trolling Victimization Factors

	Provocation Victim	Deception Victim	Distraction Victim
Provocation Victim	-	0.842	0.813
Deception Victim	0.842	-	0.765
Distraction Victim	0.813	0.765	-

Multicollinearity

Finally, multicollinearity was assessed for the measurement items of both constructs by determining their variance inflation factors (VIFs). The VIFs for all items in both constructs were below the stringent threshold value of 3.30 (Diamantopoulos & Siguaw, 2006; Kock, 2015), which means that both the trolling behavior and trolling victimization models did not have issues of multicollinearity.

Known-Group Validity

Known-group validity was also tested in this research to provide further evidence of robustness for both scales. This validity was assessed through comparing the mean scores between two groups that should, a priori, score high or low (Pons et al., 2006) on the trolling behavior and victimization constructs. For both scales, the mean scores were computed for each factor, and Welch’s t-tests were conducted (due to unequal variances and sample sizes between the groups) to determine if the difference in scores were statistically significant.

Based on prior literature, a group that should score high for both trolling behavior and victimization are frequent Internet users (Festl et al., 2013; Park et al., 2014; Savoldi & Ferraz de Abreu, 2016). Accordingly, heavy and low users of the Internet were compared to determine known-group validity. Similar to previous literature, a heavy Internet user was defined as someone who uses the Internet once or more a day (Savoldi & Ferraz de Abreu, 2016).

The results determined that the differences in the mean scores between heavy Internet users and lighter Internet users were significant ($p < 0.001$) for all trolling behavior and victimization dimensions of distraction, deception, and provocation. However, the results were not in the expected direction. For both trolling behavior and trolling victimization (see Tables 19 and 20), the mean scores were significantly higher for lighter Internet users compared to heavy Internet users. Despite the interesting direction of the results, known-group validity in terms of Internet usage was still found, albeit rather surprisingly. It is important to note that this is not an issue specific to the developed scales, because the same results were found for the GAIT scale as well. The following sections will discuss the potential explanations for the surprising results.

Regarding trolling victimization, the significant differences in the scores may be explained by the fact that trolls are often deceptive in nature, and their deception is much likelier to succeed with less experienced Internet users who are less knowledgeable and savvy about the nuances of online culture and expected behaviors compared to a heavy Internet user. Therefore, one could argue that lighter Internet users are more likely to be victims of trolling behavior, because trolls seek them out as targets.

Regarding trolling behavior, the significant differences in the scores may be explained by the fact that heavy Internet users may spend much time in certain online platforms and communities, which means they have a sense of status and membership to maintain in those spaces. Consequently, engaging in trolling behavior would be a significant risk because it could lead to them losing their standing and reputation in the online communities where they have

invested a significant portion of their time and effort. However, for lighter Internet users, they would have much less to lose in the same situation, which means, relative to heavy Internet users, they may be likelier to engage in trolling behavior.

Table 19: Mean Scores between Heavy and Light Internet Users for Trolling Behaviors

	Distraction Scores	Deception Scores	Provocation Scores
Light Internet Users	3.79	3.64	3.64
Heavy Internet Users	3.02	2.97	3.02

Table 20: Mean Scores between Heavy and Light Internet Users for Trolling Victimization

	Distraction Victim Scores	Deception Victim Scores	Provocation Victim Scores
Light Internet Users	3.98	3.89	3.92
Heavy Internet Users	3.36	3.34	3.41

Predictive Validity

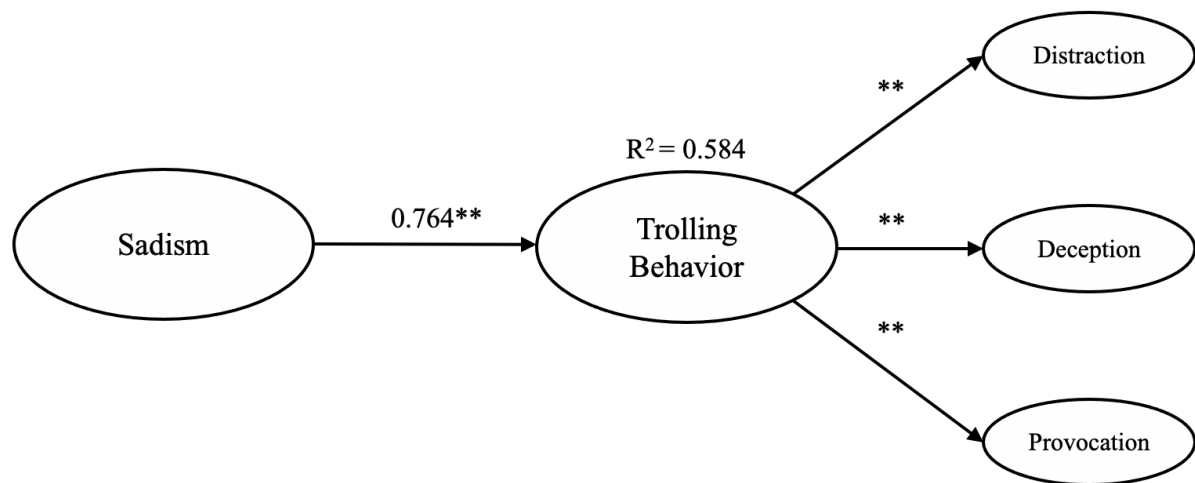
Sadism

To conclude Study 2, the predictive validity of the trolling behavior and trolling victimization constructs was investigated. To test predictive validity for trolling behavior, a 7-item scale of sadism developed by Paulhus et al. (2021) was included in the survey (see Appendix I). Sadism was included as a measure, because there is a fair amount of evidence (e.g., Buckels et al., 2014; Craker & March, 2016) demonstrating that sadism predicts trolling behavior in individuals. Accordingly, this leads to the following hypothesis:

H1: Sadism positively predicts trolling behavior.

To determine the predictive validity of the constructs, second-order models were used in the validation tests. For trolling behavior, SEM was used to test the causal relationship between sadism and trolling behavior. The SEM model found a significant path from sadism to trolling behavior with a standardized coefficient of 0.764 ($p < 0.001$) and an R^2 value of 0.584 (see Figure 4). This demonstrates how a moderate portion of the variation in trolling behavior is explained by sadism. Furthermore, the fit statistics for the SEM model were very good ($\chi^2/df = 1.773$, $p = 0.000$, RMSEA = 0.050, CFI = 0.959, TLI = 0.954, SRMR = 0.0463) and demonstrated a good-fitting model. According to these results, H1 was supported, and predictive validity was confirmed.

Figure 4: Predictive Validity Model for Trolling Behavior



* $p < 0.05$; ** $p < 0.001$

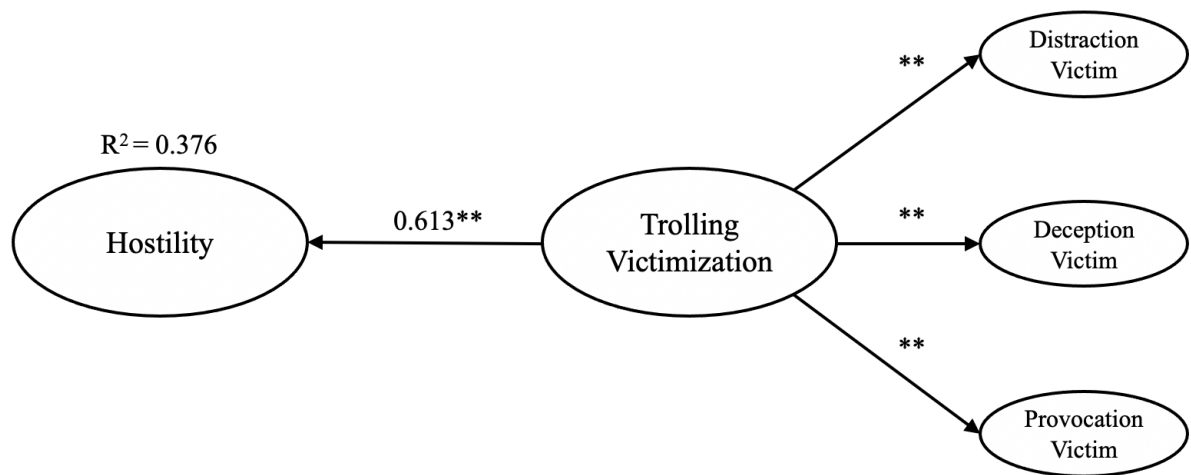
Hostility

To test predictive validity for trolling victimization, a 9-item scale of hostility adapted from Anderson and Carnagey (2009) was included in the survey (see Appendix J). Feelings of hostility was included as a measure because previous research has found that exposure to incivility online induces feelings of hostility in individuals (Gervais, 2015; Rösner et al., 2016). We believe this applies to a trolling context as well, which leads to the following hypothesis:

H2: Trolling victimization positively predicts feelings of hostility.

SEM was also used to test the causal relationship between trolling victimization and feelings of hostility. The SEM model found a significant path from trolling victimization to feelings of hostility with a standardized coefficient of 0.613 ($p < 0.001$) and an R^2 value of 0.376 (see Figure 5). Furthermore, the fit statistics for the SEM model were good ($\chi^2/df = 2.207$, $p = 0.000$, RMSEA = 0.059, CFI = 0.931, TLI = 0.924, SRMR = 0.0466) and demonstrated a fairly good-fitting model. According to these results, H2 was supported, and predictive validity was confirmed.

Figure 5: Predictive Validity Model for Trolling Victimization



* $p < 0.05$; ** $p < 0.001$

Study 3

The purpose of Study 3 was to further assess the discriminant, predictive, and concurrent validity of the trolling behavior and trolling victimization scales.

Sample

Similar to the previous studies, a short qualifier survey was created to ensure that respondents have actually engaged in trolling behavior or been victims of trolling behavior. In total, 342 responses were collected for the trolling survey, and 302 responses were collected for the victimization survey. The surveys were developed on the Qualtrics platform. The respondents were from either the United States or Canada, and they were recruited from Amazon Mechanical Turk. Respondents were compensated \$1.15 for completing both the qualifier and main survey.

After cleaning the data for careless responding in the same manner as the previous studies, 255 responses were retained for the trolling survey, and 221 responses were retained for the victimization survey. The demographic characteristics of both sets of respondents are summarized in Tables 21 and 22.

Table 21: Demographic Characteristics of Study 3 Respondents (Trolling Behavior)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	3	1.2%
United States	252	98.8%
Biological Gender		
Male	137	53.7%
Female	117	45.9%
Non-binary/Prefer to self-describe	1	0.4%
Age		
18-24 years old	12	4.7%
25-34	133	52.2%
35-44	71	27.8%
45-54	20	7.8%
55-64	15	5.9%
+65	4	1.6%

Table 22: Demographic Characteristics of Study 3 Respondents (Trolling Victimization)

Demographic Characteristics	Frequency	Percentage
Country of Residence		
Canada	8	3.6%
United States	213	96.4%
Biological Gender		
Male	100	45.2%
Female	118	53.4%
Non-binary/Prefer to self-describe	3	1.4%
Age		
18-24 years old	17	7.7%
25-34	109	49.3%
35-44	42	19.0%
45-54	36	16.3%
55-64	16	7.2%
+65	1	0.5%

Discriminant Validity

The first aim of Study 3 was to provide evidence of the discriminant validity of the trolling behavior and trolling victimization scales by showing that they are empirically distinct from similar established scales. As mentioned earlier, trolling has often been viewed as a catch-

all term for many negative online behaviors such as cyberbullying (Jussinoja, 2018; Marwick & Lewis, 2017; Sun & Fichman, 2019), which could lead to some confusion for practitioners and researchers alike. In fact, trolling behavior has been viewed both as a form of cyberbullying (Griffiths, 2014), and as a main construct where cyberbullying is an underlying factor (Miller, 2012). However, the literature on trolling and cyberbullying (see *Trolling vs. Cyberbullying* section) makes it clear that the two constructs are distinct from one another in terms of how they are directed at individuals as well as their underlying dimensions. Accordingly, this leads to the following hypotheses:

H3: Trolling behavior is empirically distinct from cyberbullying behavior.

H4: Trolling victimization is empirically distinct from cyberbullying victimization.

To test these hypotheses and avoid misuse of the constructs in future research, discriminant validity was assessed by testing a 9-item cyberbullying behavior and victimization scale (see Appendices I and J) developed by Patchin and Hinduja (2015) against the scales of trolling behavior and victimization developed in this research.

Findings

Similar to the methods of previous research (e.g., Frankowski et al., 2016; Guèvremont, 2019), discriminant validity was assessed by conducting a correlation analysis, a chi-square difference test as well as a CFI difference test (Shaffer et al., 2016). According to previous literature (Kenny, 2016; van Mierlo et al., 2009), correlations of 0.85 or higher indicate a definite lack of discriminant validity. None of the factors for both trolling behavior and trolling victimization exceeded the threshold value of 0.85 (see Tables 23 and 24), which provided initial evidence of discriminant validity.

Table 23: Discriminant Validity between Cyberbullying and Trolling Behavior Factors

	Cyberbullying
Distraction	0.803**
Deception	0.736**
Provocation	0.732**

* $p < 0.05$; ** $p < 0.001$

Table 24: Discriminant Validity between Cyberbullying and Trolling Victimization Factors

	Cyberbullying Victim
Distraction Victim	0.590**
Deception Victim	0.316**
Provocation Victim	0.371**

* $p < 0.05$; ** $p < 0.001$

Next, for the chi-square and CFI difference tests, one- and three-factor models of trolling behavior and trolling victimization were tested against one-factor models of cyberbullying behavior and cyberbullying victimization. According to Shaffer et al. (2016), when the CFI difference between a constrained model and unconstrained models is greater than

0.002, then this provides evidence of discriminant validity. For the one-factor models of trolling behavior and trolling victimization, the model comparisons showed that constraining covariance between the constructs to 1 led to significantly reduced fit for both trolling behavior ($\Delta\chi^2 = 19.093$, $\Delta df = 1$, $\Delta CFI = 0.004$, and $p = 0.000$), and trolling victimization ($\Delta\chi^2 = 28.276$, $\Delta df = 1$, $\Delta CFI = 0.006$, and $p = 0.000$). For the three-factors models, the same results occurred for both trolling behavior ($\Delta\chi^2 = 39.685$, $\Delta df = 6$, $\Delta CFI = 0.007$, and $p = 0.000$) and trolling victimization ($\Delta\chi^2 = 78.271$, $\Delta df = 6$, $\Delta CFI = 0.017$, and $p = 0.000$). This provides further evidence of discriminant validity between the trolling behavior and cyberbullying constructs as well as the trolling victimization and cyberbullying victimization constructs. Thus, H3 and H4 were supported.

Validity in a Nomological Network: Trolling Behavior

The second aim of Study 3 was to further assess the validity of trolling behavior by examining the construct as part of a larger nomological network with multiple constructs. As a note, the sadism construct (testing H1) from the predictive validity tests in Study 2 was also included in this examination.

Similar to the previous predictive validity tests from Study 2 and previous research (e.g., Busser & Shulga, 2018; Öberseder et al., 2014), validity was tested using the second-order trolling behavior models. SEM was again used to test the causal relationships among the predictor constructs and trolling behavior. Similar to the methods in previous research, the SEM models were refined by removing a few predictor items one at a time due to high modification indices or low factor loadings (Ahmad et al., 2016; Bagozzi & Yi, 1988). The validity of trolling behavior was further assessed by adding the constructs of sadism, depression, psychopathy, and agreeableness to its nomological network (see Appendix I).

Sadism

H1 was tested again with the same 7-item scale developed by Paulhus et al. (2021). The SEM model again found a significant path from sadism to trolling behavior with a standardized coefficient of 0.774 ($p < 0.001$) and an R^2 value of 0.599 (see Figure 6). Furthermore, the fit statistics for the SEM model were very good ($\chi^2/df = 1.730$, $p = 0.000$, RMSEA = 0.054, CFI = 0.953, TLI = 0.946, SRMR = 0.0514) and showed a good-fitting model. According to these results, H1 was further supported, which demonstrates how sadism is a significant predictor of trolling behavior.

Depression

According to prior literature, depression has been found to predict trolling and cyberbullying behavior in individuals (Guo, 2016; Hong & Cheng, 2018), which led to its inclusion in this nomological network. Depression was measured with a 7-item scale developed by Lovibond and Lovibond (1995). Accordingly, this led to the following hypothesis:

H5: Depression positively predicts trolling behavior.

The SEM model found a significant path from depression to trolling behavior with a standardized coefficient of 0.738 ($p < 0.001$) and an R^2 value of 0.544 (see Figure 6). The fit statistics for the SEM model were excellent ($\chi^2/df = 1.451$, $p = 0.000$, RMSEA = 0.042, CFI = 0.974, TLI = 0.971, SRMR = 0.0345) and verified a good-fitting model. According to these

results, H5 was supported, which illustrates how depression is a significant predictor of trolling behavior.

Psychopathy

Next, psychopathy was included in the network because it has been found to be a predictor of trolling behavior in previous research (Buckels et al., 2014; Craker & March, 2016). Psychopathy was measured with a 4-item scale developed by Jonason and Webster (2010). Accordingly, this leads to the following hypothesis:

H6: Psychopathy positively predicts trolling behavior.

The SEM model was run and found a significant path from psychopathy to trolling behavior with a standardized coefficient of 0.887 ($p < 0.001$) and an R^2 value of 0.787 (see Figure 6). Once again, the fit statistics for the SEM model were excellent ($\chi^2/df = 1.637$, $p = 0.000$, RMSEA = 0.050, CFI = 0.970, TLI = 0.965, SRMR = 0.0366) and confirmed a good-fitting model. According to these results, H6 was confirmed, which shows how psychopathy is a significant predictor of trolling behavior.

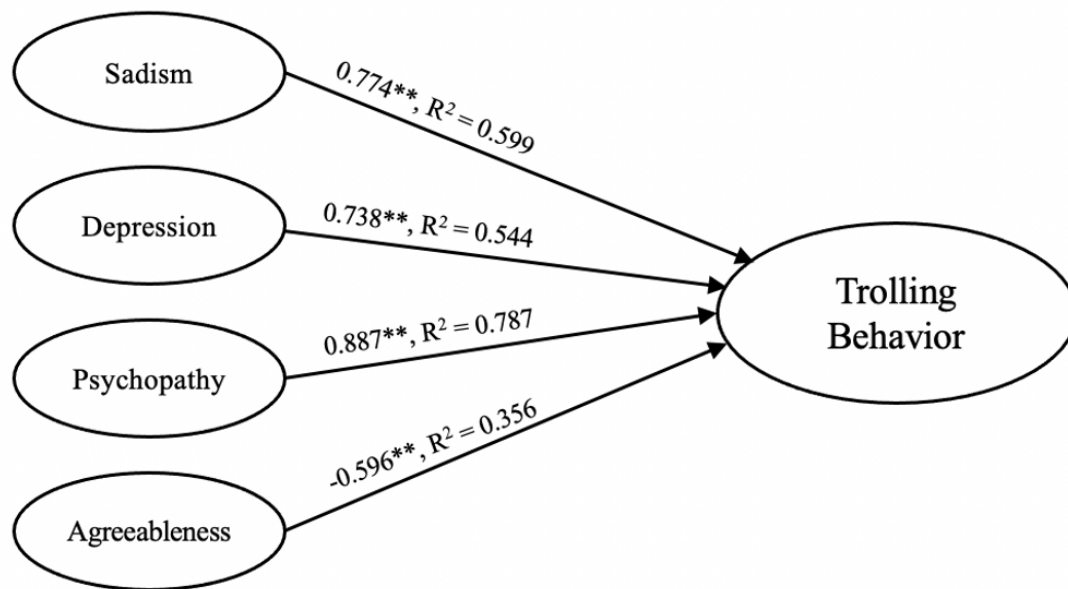
Agreeableness

Finally, the Big Five personality factor of agreeableness was included due to its negative relationship to cyberbullying and trolling behavior in individuals (Balakrishnan et al., 2019; Hong & Cheng, 2018; van Geel et al., 2017). Agreeableness was measured with a 4-item scale of agreeableness from the Mini-IPIP scale developed by Donnellan et al. (2006). Accordingly, this led to the following hypothesis:

H7: Agreeableness negatively predicts trolling behavior.

The SEM model found a significant and negative path from agreeableness to trolling behavior with a standardized coefficient of -0.596 ($p < 0.001$) and an R^2 value of 0.356 (see Figure 6). Once more, the fit statistics for the SEM model were excellent ($\chi^2/df = 1.247$, $p = 0.029$, RMSEA = 0.031, CFI = 0.988, TLI = 0.985, SRMR = 0.0404) and established a good-fitting model. According to these results, H7 was supported, which confirms how agreeableness is a significant negative predictor of trolling behavior.

Figure 6: SEM Model for the Nomological Network of Trolling Behavior



* $p < 0.05$; ** $p < 0.001$

Correlation Analyses

Finally, correlation analyses were performed to test all the relationships among trolling behavior and the predictor constructs (see Table 25). All relationships were found to be significant ($p < 0.001$). According to the results from the SEM models and correlation analyses, H1, H5, H6, and H7 were supported, which provides more evidence of how trolling behavior is significantly associated with other relevant and established constructs as part of a larger nomological network.

Table 25: Correlation Analysis for Trolling Behavior Constructs

	Sadism	Depression	Psychopathy	Agreeableness
Distraction	0.640**	0.603**	0.731**	-0.429**
Deception	0.556**	0.576**	0.678**	-0.443**
Provocation	0.638**	0.620**	0.690**	-0.452**

* $p < 0.05$; ** $p < 0.001$

Validity in a Nomological Network: Trolling Victimization

The third aim of Study 3 was to further assess the validity of trolling victimization by examining the construct as part of a larger nomological network with multiple constructs. As a note, the hostility construct (testing H2) from the predictive validity tests in Study 2 was also included in this assessment.

The methods used to assess the validity of the trolling behavior construct in the previous section were also used to determine the validity of the trolling victimization construct. In other words, SEM was used to test the causal relationships among the predictor and consequence constructs, and trolling victimization. The victimization SEM model was also refined by removing predictor items one at a time due to high modification indices or low factor loadings

(Ahmad et al., 2016; Bagozzi & Yi, 1988). To start, the validity of trolling victimization was further assessed by adding the hostility, depression, willingness to divulge personal information, and introversion constructs to its nomological network (see Appendix J).

Hostility

H2 was tested again with the same 9-item scale used in Study 2 developed by Anderson and Carnagey (2009). Just like in Study 2, the SEM model found a significant path from trolling victimization to hostility with a standardized coefficient of 0.541 ($p < 0.001$) and an R^2 value of 0.293 (see Figure 7). In addition, the fit statistics for the SEM model were good ($\chi^2/df = 1.756$, $p = 0.000$, RMSEA = 0.059, CFI = 0.943, TLI = 0.936, SRMR = 0.0527) and revealed a good-fitting model. According to these results, H2 was further supported, which demonstrates how trolling victimization is a significant predictor of hostility.

Depression

Depression was also included in the nomological network due to how it has been found to be a predictor to trolling and cyberbullying victimization in previous research (Guo, 2016; Hong & Cheng, 2018). Depression was measured with the same 7-item scale developed by Lovibond and Lovibond (1995). Accordingly, this led to the following hypothesis:

H8: Depression positively predicts trolling victimization.

The SEM model found a significant path from depression to trolling victimization with a standardized coefficient of 0.527 ($p < 0.001$) and an R^2 value of 0.278 (see Figure 7). Also, the fit statistics for the SEM model were fairly good ($\chi^2/df = 2.040$, $p = 0.000$, RMSEA = 0.069, CFI = 0.930, TLI = 0.922, SRMR = 0.0731), which showed a good-fitting model. According to these results, H8 was supported, which proves how depression is a significant predictor of trolling victimization.

Willingness to Divulge Personal Information

Next, a willingness to divulge personal information or private matters online was included in the nomological network due to its connection to victimization in individuals (Costello et al., 2017; Hawdon et al., 2014). This construct was measured with a 6-item scale measuring an individual's willingness to provide personal information adapted from scales developed by Phelps et al. (2000) and Gupta et al. (2010). Accordingly, this led to the following hypothesis:

H9: A willingness to divulge personal information positively predicts trolling victimization.

The SEM model did not find a significant path from a willingness to divulge personal information to trolling victimization with a standardized coefficient of 0.134 ($p = 0.082$) and an R^2 value of 0.018 (see Figure 7). However, the fit statistics were fairly good ($\chi^2/df = 1.816$, $p = 0.000$, RMSEA = 0.061, CFI = 0.943, TLI = 0.935, SRMR = 0.0666). Overall, H9 was not supported, which means a willingness to divulge personal information is not a significant predictor of trolling victimization. The lack of support for H9 may be due to the fact that trolling is not necessarily directed at individuals who share fairly neutral pieces of information such as

their demographics, because that cannot be easily mocked. Instead, trolling is often directed at the assumed opinions and/or feelings of other individuals or communities (Coles & West 2016a, Cruz et al., 2018; Sanfilippo et al., 2018), because that is what often elicits a reaction from individuals. This may explain the lack of support for H9.

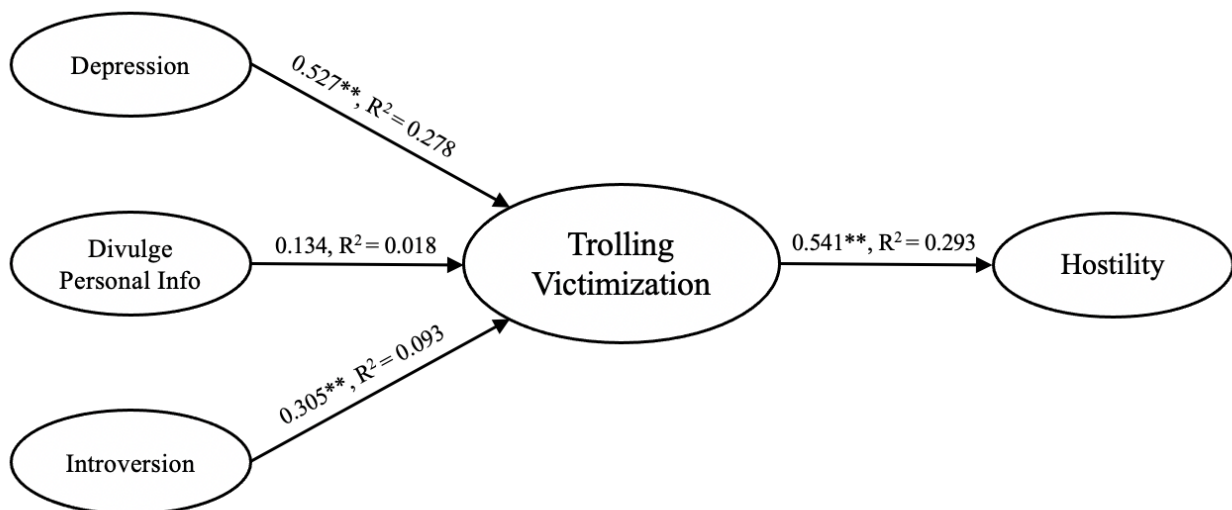
Introversion

Finally, social introversion was added due to how it has been perceived as an antecedent of victimization in online (Festl & Quandt, 2013) and offline contexts (Glasø et al., 2007; Olweus, 2014). Introversion was measured with a 4-item scale of introversion from the Mini-IPIP scale developed by Donnellan et al. (2006). Accordingly, this led to the following hypothesis:

H10: Social introversion positively predicts trolling victimization.

The SEM model found a significant path from introversion to trolling victimization with a standardized coefficient of 0.305 ($p < 0.001$) and an R^2 value of 0.093 (see Figure 7). Also, the fit statistics for the SEM model were fairly good ($\chi^2/df = 1.906$, $p = 0.000$, RMSEA = 0.064, CFI = 0.939, TLI = 0.929, SRMR = 0.0749), which confirms a good-fitting model. According to these results, H10 was supported, which confirms how introversion is a significant predictor of trolling victimization.

Figure 7: SEM Model for the Nomological Network of Trolling Victimization



* $p < 0.05$; ** $p < 0.001$

Correlation Analyses

Finally, correlation analyses were conducted to test all the relationships among trolling victimization and the predictor and consequence constructs (see Table 26). All the relationships apart from the ‘divulge personal information → deception’ and the ‘divulge personal information → provocation’ relationships were found to be significant ($p < 0.001$). According to the results from the SEM models and correlation analyses, H2, H8, and H10 were supported,

which demonstrates how trolling victimization is significantly associated with other relevant and established constructs as part of a larger nomological network.

Table 26: Correlation Analysis for Trolling Victimization Constructs

	Hostility	Depression	Divulge Personal Information	Introversion
Distraction Victim	0.462**	0.582**	0.257**	0.239**
Deception Victim	0.371**	0.260**	0.030	0.213**
Provocation Victim	0.438**	0.363**	0.045	0.284**

* $p < 0.05$; ** $p < 0.001$

Concurrent Validity

The final objective of Study 3 was to assess the concurrent validity of both the trolling behavior and trolling victimization scales. This was accomplished by comparing the trolling scales developed in this study (Haverila, 2023) against the established Global Assessment of Internet Trolling (GAIT) scale developed by Buckels et al. (2014) (see Appendices I and J). The predictor and consequence constructs from the previous sections were included to determine which models had better fit statistics. It is important to note that the willingness to divulge personal information construct was not included, because H9 was not confirmed in the prior analysis. Also, the models were not directly compared against one another (i.e., by themselves with no predictor or consequence constructs) due to the low degrees of freedom present in the 4-item GAIT scale, which led to inflated fit statistics.

Findings

SEM tests were run and the fit statistics of both scales were assessed to determine which model had better fit. Similar to the methods used in previous research, the SEM model was refined by removing predictor items one at a time due to high modification indices or low factor loadings (Ahmad et al., 2016; Bagozzi & Yi, 1988). The full list of comparison results, which includes fit statistics and path coefficients for trolling behavior and trolling victimization can be found in Tables 27 and 28 respectively. For trolling behavior, all relationships had better fit statistics for the Haverila (2023) model compared to the GAIT model (Buckels et al., 2014). For trolling victimization, most relationships had better fit statistics for the Haverila (2023) model compared to the GAIT model (Buckels et al., 2014) apart from the CFI, TLI and SRMR statistics in the ‘depression → victimization’ relationship. Furthermore, unlike the GAIT models, the fit statistics for each relationship in the Haverila (2023) trolling behavior and trolling victimization models met the minimum threshold values determined by previous research (e.g., Hair et al., 2010), which provides evidence of good-fitting models. The path coefficients were all significant for both constructs. Finally, the correlations between the Haverila (2023) scales and the GAIT scales were significant for both trolling behavior (0.751) and victimization (0.631). Overall, these results provide sound evidence of concurrent validity for both constructs and further support for H1, H2, H5, H6, H7, H8, and H10.

Table 27: Comparison between Trolling Behavior and GAIT Scales

	Model	χ^2/df	RMSEA	CFI	TLI	SRMR	Path coefficient
Sadism → Trolling	<i>Haverila (2023)</i>	1.730	0.054	0.953	0.946	0.0514	0.774**
	<i>GAIT</i>	2.667	0.081	0.938	0.920	0.0472	0.835**
Depression → Trolling	<i>Haverila (2023)</i>	1.451	0.042	0.974	0.971	0.0345	0.738**
	<i>GAIT</i>	2.543	0.078	0.958	0.946	0.0369	0.734**
Psychopathy → Trolling	<i>Haverila (2023)</i>	1.637	0.050	0.970	0.965	0.0366	0.887**
	<i>GAIT</i>	3.297	0.095	0.961	0.942	0.0379	0.878**
Agreeableness → Trolling	<i>Haverila (2023)</i>	1.247	0.031	0.988	0.985	0.0404	-0.596**
	<i>GAIT</i>	2.934	0.087	0.958	0.932	0.0609	-0.453**

* $p < 0.05$; ** $p < 0.001$

Table 28: Comparison between Trolling Victimization and GAIT Victimization Scales

	Model	χ^2/df	RMSEA	CFI	TLI	SRMR	Path coefficient
Victimization → Hostility	<i>Haverila (2023)</i>	1.756	0.059	0.943	0.936	0.0527	0.541**
	<i>GAIT Victim</i>	2.682	0.087	0.943	0.927	0.0826	0.271**
Depression → Victimization	<i>Haverila (2023)</i>	2.040	0.069	0.930	0.922	0.0731	0.527**
	<i>GAIT Victim</i>	2.375	0.079	0.965	0.956	0.0340	0.749**
Introversion → Victimization	<i>Haverila (2023)</i>	1.906	0.064	0.939	0.929	0.0749	0.305**
	<i>GAIT Victim</i>	5.251	0.139	0.899	0.837	0.1289	0.353**

* $p < 0.05$; ** $p < 0.001$

Discussion

Results

Overall, the pre-tests and three studies in this research aimed to accurately conceptualize the constructs of trolling behavior and trolling victimization. In the pre-test and expert review, the survey measurement items were initially generated to provide a comprehensive albeit over-inclusive description of the trolling behavior and trolling victimization scales. In Study 1, the scales were then refined and purified with an EFA, which initially revealed the multidimensionality of both the trolling behavior and trolling victimization constructs. In Studies 2 and 3, the multidimensionality of both constructs were tested and confirmed with several tests of reliability and validity accomplished through SEM. Finally, trolling behavior and trolling victimization were examined in a nomological network with multiple established constructs, which provided further evidence of psychometric validation for both constructs.

Implications

Theoretical Implications

This research contributes to trolling literature by providing a more comprehensive view of trolling behavior and trolling victimization. This was accomplished by discovering the multidimensionality of both constructs where distraction, deception, and provocation were the main underlying factors in both cases. Furthermore, the domain of trolling as well as its boundaries have been more suitably delineated through the development of this scale. Evidently, there is still much more to learn about trolling as a concept; however, this research is a vital step in the right direction that further solidifies the foundation of the trolling construct.

This research also shows how trolling can be practiced in at least three specific ways. First, trolling can come in the form of deliberately disrupting online discussions and creating a general sense of confusion. Second, trolling can be achieved by deceiving individuals through lying and creating fake profiles. Finally, trolling can also have a provocative side where the focus is placed on creating conflict and generating emotional reactions from others. These findings are in line with the findings of previous literature where trolling behavior has been found to be a highly context-dependent behavior that can be practiced in numerous ways across communities and platforms (Dyrel, 2016; Fichman & Peters, 2019; Sanfilippo et al., 2018; Sun & Fichman, 2019). Crucially, this is the first research that provides empirical evidence of the multidimensional nature of trolling behavior and victimization. While there is much more work to be done when it comes to understanding trolling behavior and victimization, this research has done some solid groundwork that should hopefully lead to further research on the constructs.

Another key contribution of this research is that a scale of trolling victimization has been developed. Previous trolling research has mainly focused on defining and conceptualizing trolling behavior itself, but little attention has been placed on trolling victimization. It is vital to investigate how trolling can affect targeted individuals, communities, and other relevant groups, because they are the ones feeling the impact of the behavior. Accordingly, a scale of trolling victimization was included in this research to determine the wide-ranging effects that trolling can have a number of different actors. This scale was a crucial addition, because it contributes to a more well-rounded view of the trolling construct where multiple actors, i.e., the troll and the target of the behavior, are being examined. This research also discovered how

trolling victimization is measured slightly differently compared to trolling behavior. This is an intriguing finding, because it demonstrates how victims and perpetrators have different views on what constitutes trolling behavior. For instance, the item of “*creating false reviews*” was found to be a valid and reliable item for trolling victimization, but not for trolling behavior. This may be explained by the fact that being a victim of false reviews can have a relatively significant and noticeable impact on the victim of the behavior. However, for a troll, creating a false review will not generate an instantaneous and noticeable reaction from the target, which is what the troll typically desires, which may explain why the item was not included in the final trolling scale. Overall, these results contribute to the findings of previous research (e.g., Sanfilippo et al., 2018) where trolling and its victimization have been found to be highly relational constructs that can be perceived very differently based on the individuals and the situation under investigation.

Finally, the relationships between trolling behavior and multiple constructs were examined to determine the validity of the scale in a nomological network; the same was done for trolling victimization as well. Overall, the validity tests supported the proposed hypotheses where sadism, psychopathy, and depression were found to be positive predictors of trolling behavior, and agreeableness was found to be a negative predictor of trolling behavior. For trolling victimization, depression and social introversion were found to be positive predictors of trolling victimization. Finally, feelings of hostility were confirmed as an outcome of trolling victimization. In the study results, all but one of the results were consistent with the findings of previous research, which examined the relationships of similar negative constructs such as online incivility, and cyberbullying perpetration and victimization (e.g., Guo, 2016; Hong & Cheng, 2018; van Geel et al., 2017). Also, when comparing models, the trolling behavior and trolling victimization scales were found to fit better than the established GAIT scale in most instances. Overall, the fact that predictive and concurrent validity was well established for both the trolling behavior and trolling victimization scales illustrates the robustness of both constructs.

Practical Implications

Developing scales for trolling behavior and trolling victimization can provide some benefits from a practical standpoint as well. When managers and other practitioners better understand the trolling behavior and victimization constructs, then they are better equipped to identify and manage the behavior when it occurs on their online platforms and communities. Managers of online communities can devise prevention strategies to deal with trolling behavior and ideally mitigate its effects on both the individual members as well as the community as a whole. Furthermore, managers can also create coping strategies to aid the targets and victims of the behavior to minimize the potential negative effects that it could have on the customer-brand and customer-community relationships. This is a vital finding, because previous research has found that trolling can have negative impacts on the individual-, community-, and brand-level (e.g., Coles & West, 2016a; Demers et al., 2021; Golf-Papez & Veer, 2017). Overall, developing the two scales is a net positive for all relevant actors, because managers are better equipped to deal with the troll as well as the victims who have been affected.

Finally, the scales for trolling behavior and trolling victimization can be adapted and used on a broader level, which can be valuable for practitioners. More specifically, the measurement items can be used to test trolling behavior on multiple levels where the trolling can be directed at individual members, employees, the community as a whole, or even the brand itself. Overall, the scales can be adapted to conduct trolling research on a broader scale in novel online settings, and further build on the findings of previous research that has examined the

wider and more far-reaching impact of trolling behavior (e.g., Coles & West, 2016a; Demsar et al., 2021; Golf-Papez & Veer, 2017).

Future Research Avenues and Limitations

This study contributes to trolling literature by creating an empirically reliable and valid scale of trolling behavior and trolling victimization. However, some limitations exist in this research, which opens up avenues for future research. Firstly, the responses were only gathered from a North American sample where the majority of the responses came from the United States. This is problematic because the generalizability of the results is limited and cannot necessarily be applied across populations. Accordingly, a cross-cultural examination of the constructs that include respondents from a wider array of countries could solve this issue.

Secondly, there will always be some methodological and epistemological challenges when measuring and self-reporting negative behaviors such as trolling. These issues include but are not limited to finding willing respondents (Shachaf & Hara, 2010), self-reported honesty issues (Buckels et al., 2018), and social desirability biases (Cheng et al., 2017; Frankowski et al., 2016; Podsakoff et al., 2003). The impact of these issues was minimized by following proper scale development procedures such as including a qualifier survey, guaranteeing that responses were confidential, allowing the respondents to complete the surveys at their own pace and time, and applying post-hoc statistical remedies to check for common method bias. However, future research could include some social desirability measures to enhance the trolling behavior scale's robustness, reliability, and validity.

Thirdly, while the trolling behavior and trolling victimization scales were found to be multidimensional, there may be more to discover regarding the dimensionality of the constructs. For instance, in the content analysis of 212 research papers, the factors of maliciousness, aggression, and antagonism were mentioned fairly frequently. Furthermore, previous research has also examined the humorous nature of trolling (Sanfilippo et al., 2018) as well as its potential to be positive in certain contexts (Cruz et al., 2018). Overall, this demonstrates that there is potentially more to discover when it comes to measuring trolling behavior and trolling victimization in the future.

Fourthly, while the scales' reliability and validity were assessed fairly comprehensively, a few more validity tests could have been included to strengthen the robustness of the scales. Specifically, future research could assess test-retest reliability to determine the stability of the scales over time. Also, concurrent validity could be further tested with additional scales such as the iTroll questionnaire developed by Buckels et al. (2018) or the "Conceptions of Online Trolls" scale developed by Maltby et al. (2016) to provide further evidence of robustness for the scales developed in this research.

Finally, this research focused on trolling behavior and victimization from an individual perspective. It would be worthwhile to gain a more marketing-oriented perspective on trolling and its victimization. This is especially true when extant literature has begun to discuss the concept of brand trolling where the impact of trolling can extend beyond the individual to employees and brands themselves (Demsar et al., 2021). In addition, previous research has mentioned the possibility of brands engaging in trolling themselves (e.g., Behl & Bhutani, 2022; Golf-Papez & Veer, 2022). Accordingly, to gain a better understanding of trolling from a marketing viewpoint, a brand trolling behavior and victimization scale could be developed where responses are specifically gathered from devoted brand community members and employees such as brand community managers.

Conclusions

However, even with these aforementioned limitations, this research further advanced the knowledge on trolling behavior and its victimization by developing reliable and valid psychometric instruments to measure both constructs. Overall, the results of this research build on the initial conceptual findings of *Essay I* and further solidify the foundation of the trolling construct. While there is much more work to be done with regards to issues of generalizability, dimensionality, and validity of the instruments, it is the hope of the author that these scales will serve as a foundation that will allow scholars to be better prepared to research trolling behavior and victimization. Consequently, this should aid in the identification and management of trolling behavior thereby making future online communities more harmonious and valuable for members and organizations alike.

Dissertation Concluding Remarks

The growing prominence and use of online technologies and social media platforms (De' et al., 2020; United Nations, 2023) has made online communication, activities, and interactions both a prevalent and essential part of life (McClain et al., 2021). However, with this increasing use of technology, it means individuals are likelier to observe and engage in both positive and negative behaviors in online environments. This dissertation specifically focused on negative trolling behaviors and their victimization due to its rising frequency in online environments (Golf-Papez & Veer, 2022; YouGov, 2014), and its potential to seriously impact individuals, brands and organizations (Demsar et al., 2021; Golf-Papez & Veer, 2017, 2022). Despite the growing literature on trolling, some key research gaps still exist. First, research on trolling from a marketing perspective has been relatively scant and often narrow in scope with rather disjointed findings on a collective level, which indicates a lack of a proper foundation. Second, trolling generally lacks conceptual clarity and has not been comprehensively conceptualized in prior research, which results in confusion among practitioners and scholars. This dissertation addressed these gaps in the literature with two essays.

Essay I focused on developing a more solid foundation for trolling research in the marketing domain. As a first step, this was achieved through introducing the concept of brand trolling which shifted the context from trolling taking place in any online environment directed at any individual to trolling occurring in online brand communities directed at community members and other organization- and brand-related actors. This is a crucial shift in perspective, because it begins to suggest how trolling can also be impactful in the marketing sector. Next, the foundation was further solidified by broadening the existing scope of brand trolling research by generating relevant research propositions and questions with a more holistic viewpoint in mind. More specifically, these propositions aimed to move beyond the typical dyadic troll-target relationship and instead emphasize the role of other actors such as bystanders and employees who are often involved and victimized in trolling situations. Also, the propositions stressed how attention should be placed on the antecedents and consequences of trolling-related constructs. This approach is significant, because assessing the empirical relationships of novel constructs demonstrates their robustness as concepts moving forward. Overall, this conceptual research highlighted how brand trolling is a novel construct that has plenty of potential to be investigated from a multitude of different perspectives on several levels.

Essay II concentrated on developing a measurement framework that makes it possible for the foundational and empirical research propositions and questions developed in *Essay I* to be answered. More specifically, this essay aimed to develop psychometric measurement scales for both trolling behavior and trolling victimization. Proper scale development practices and steps (e.g., Churchill, 1979; Gerbing & Anderson, 1988; Worthington & Whittaker, 2006) were followed to develop the scales and ensure their reliability and validity. First, an EFA was completed to purify the scale and a CFA was conducted to further validate the scale. Next, the predictive and nomological validity was assessed to determine how trolling behavior and victimization interact with established constructs that they are hypothesized to be related to. Following this, the constructs were tested against the related construct of cyberbullying to determine their discriminant validity. Finally, the concurrent validity of both constructs was assessed by testing them against the established Global Assessment of Internet Trolling (GAIT) scale. All of these rigorous procedures confirmed the reliability, validity, and multidimensionality of both the 15-item trolling behavior and 16-item trolling victimization scales. On balance, the findings in *Essay II* further substantiate the foundation of trolling-related research in the marketing context by developing multiple instruments that are essential to conduct research, and answer the questions and propositions set forth in *Essay I*.

Essay I produced three key theoretical contributions. First, the essay broadened the knowledge of trolling and brand trolling where the focus shifted from micro-level relationships between the troll and the target to broader meso- and macro-level relationships that included numerous actors such as community members, employees, the communities themselves, brands and even organizations. Second, this shift in perspective opens up countless possibilities to research novel relationships as well as previously determined relationships from prior research to see if they also apply in the brand trolling context. Finally, the third contribution illustrates how the brand community setting opens up a wide range of research possibilities due to their highly diverse and contextual nature where brand trolling can greatly vary due to the many differences that can exist among brand communities. Overall, these contributions highlight how brand trolling is a highly extensive and diverse topic that can be investigated from many different angles.

Essay II also generated three theoretical contributions. First, the essay contributes to a more well-rounded view of the trolling behavior and trolling victimization constructs through the development of robust multidimensional scales. Second, this essay confirms that trolling can be carried out in at least three unique ways through confirming the three dimensions of distraction, deception, and provocation. This finding provides further evidence of the highly contextual and adaptive nature of trolling. Finally, this research investigated the trolling victimization construct, which is a crucial step in the right direction, because it opens the door for research that addresses the potential serious consequences that trolling can have on individuals, brands, and organizations (Demsar et al., 2021; Golf-Papez & Veer, 2017, 2022).

Essay I has some implications that are relevant on a practical level. First, establishing the brand trolling construct makes it easier for community employees and managers to identify and manage the behavior in their respective communities. Next, the introduction of brand trolling inevitably leads to managers developing management strategies to deal with the behavior. However, managers must set the correct expectations about community conduct and behavior so that the communities can continue to serve their original purpose. Finally, the research propositions highlighted how there is an amplified role for bystanders and other community members in the brand trolling context. This is because they are partially responsible for ensuring that the conduct in the community is appropriate through their ability to report behavior that goes against the guidelines. Consequently, managers should foster close communities to encourage this protective behavior so that the communities can continue to flourish.

Finally, the findings from *Essay II* also generate two practical implications. First, generating scales that specify the exact practices and behaviors of trolling provides some much needed clarity for managers and practitioners who constantly have to deal with the behavior. More specifically, having objective measures in place makes it easier for managers to identify and manage trolling when it occurs. With this knowledge, managers can establish prevention and coping strategies to hopefully minimize the effects of trolling in their communities. The second practical implication relates to how these measurement scales can be adapted and utilized for different purposes and situations. For instance, the items could be adapted to several online areas such as brand communities, social media platforms, forums, online gaming, comment sections, and much more. All in all, the versatility of these scales can be used to assess the far-reaching impact of trolling and help managers deal with the behavior in a multitude of different online contexts.

Overall, *Essays I* and *II* contribute to laying the groundwork for trolling in the marketing domain. Essentially, they are the crucial first steps of the research process that must be taken when developing a foundation for a novel construct. Without a solid foundation, there will continue to be research with piecemeal findings that do not significantly contribute to the knowledge of trolling behavior and victimization. As November (2004) discussed, scholars

often view research gaps as bricks used to create a wall. Yet, those bricks often end up creating a pile of bricks instead. A researcher may understand what a brick is, but the bricks will not self-organize into a wall. It seems as if there is little agreement on which bricks should be made first, because there is no foundation to the wall. There are no wall designers or architects with a clear idea of what the eventual wall should look like when it is complete. From a trolling perspective, the focus should be on building the foundation through exploratory conceptual research and scale development, which is what *Essays I* and *II* achieved.

Accordingly, scholars should now place their focus on empirical research to further contribute to the knowledge of trolling. It is the hope of the author that the findings of *Essay I* encourage more investigation into brand trolling and its surrounding relationships. Simply put, if future empirical research begins to answer the initial research propositions and questions set forth in *Essay I*, then this would be a vital step in the right direction for trolling literature. Finally, it would be ideal if the robustness of the scales developed in *Essay II* would be further enhanced in future research. For instance, applying the scales in novel cultural, national, and online contexts with responses from devoted brand community members and employees would improve their rigor. Moreover, adding new dimensions such as aggression and toxicity could also prove to be a valuable next step. Overall, these future research considerations are critical to push the knowledge of trolling forward beyond the initial foundation. The first steps have been taken, now the journey of discovery must continue.

References

- Ahmad, S., Zulkurnain, N. N. A., & Khairushalimi, F. I. (2016). Assessing the fitness of a measurement model using confirmatory factor analysis (CFA). *International Journal of Innovation and Applied Studies*, 17(1), 159-168.
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The 'nasty effect': Online incivility and risk perceptions of emerging technologies. *Journal Of Computer-Mediated Communication*, 19(3), 373–387. doi:10.1111/jcc4.12009
- Anderson, C. A., & Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: Is it competitiveness or violent content? *Journal of Experimental Social Psychology*, 45(4), 731–739. <https://doi.org/10.1016/j.jesp.2009.04.019>
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94. <https://doi.org/10.1007/bf02723327>
- Balakrishnan, V., Khan, S., Fernandez, T., & Arabnia, H. R. (2019). Cyberbullying detection on twitter using Big Five and Dark Triad features. *Personality and Individual Differences*, 141, 252–257. <https://doi.org/10.1016/j.paid.2019.01.024>
- Baldus, B. J., Voorhees, C., & Calantone, R. (2015). Online brand community engagement: Scale development and validation. *Journal of Business Research*, 68(5), 978-985.
- Behl, A., & Bhutani, C. (2022). Dark Side of Online Brand Communities: A Scoping Review of the Literature. *SSRN*, 1–49. <https://ssrn.com/abstract=4088130>
- Berendt, J., Uhrich, S., & Thompson, S. A. (2018). Marketing, get ready to rumble—how rivalry promotes distinctiveness for brands and consumers. *Journal of Business Research*, 88, 161–172. <https://doi.org/10.1016/j.jbusres.2018.03.015>
- Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science*, 29(5), 815–827. <https://doi.org/10.1287/mksc.1090.0557>
- Binns, A. (2012). Don't feed the trolls! *Journalism Practice*, 6(4), 547–562. <https://doi.org/10.1080/17512786.2011.648988>
- Bishop, J. (2013). The art of trolling law enforcement: A review and model for implementing 'flame trolling' legislation enacted in Great Britain (1981–2012). *International Review of Law, Computers & Technology*, 27(3), 301–318. <https://doi.org/10.1080/13600869.2013.796706>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and

- Behavioral Research: A Primer. *Frontiers in Public Health*, 6, 1–18.
<https://doi.org/10.3389/fpubh.2018.00149>
- Breitsohl, J. C., Jimenez, N., & Roschk, H. (2021). Investigating consumers' motives for consumer brand-cyberbullying on social media. *The Information Society*, 38(1), 1–12.
<https://doi.org/10.1080/01972243.2021.1981507>
- Breitsohl, J. C., Roschk, H., & Feyertag, C. (2018). Consumer brand bullying behaviour in online communities of service firms. In M. Bruhn & K. Hadwich (Eds.), *Service Business Development*. (pp. 1-24). Springer Gabler.
- Buckels, E. E., Trapnell, P. D., Andjelovic, T., & Paulhus, D. L. (2018). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality*, 87(2), 328-340.
- Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97-102.
- Buelga, S., Postigo, J., Martínez-Ferrer, B., Cava, M.-J., & Ortega-Barón, J. (2020). Cyberbullying among adolescents: Psychometric properties of the Cyb-AGS cyber-aggressor scale. *International Journal of Environmental Research and Public Health*, 17(9), 1–15. <https://doi.org/10.3390/ijerph17093090>
- Busser, J. A., & Shulga, L. V. (2018). Co-created value: Multidimensional scale and nomological network. *Tourism Management*, 65, 69–86.
<https://doi.org/10.1016/j.tourman.2017.09.014>
- Carpenter, S. (2017). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, 12(1), 25–44.
<https://doi.org/10.1080/19312458.2017.1396583>
- Caspi, A., & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior*, 9(1), 54–59. <https://doi.org/10.1089/cpb.2006.9.54>
- Çetin, B., Yaman, E., & Peker, A. (2011). Cyber victim and bullying scale: A study of validity and reliability. *Computers & Education*, 57(4), 2261–2271.
<https://doi.org/10.1016/j.compedu.2011.06.014>
- Chandler, J. D., & Vargo, S. L. (2011). Contextualization and value-in-context: How context frames exchange. *Marketing Theory*, 11(1), 35–49.
<https://doi.org/10.1177/1470593110393713>
- Chen, Y. (2018). “Being a butt while on the internet”: Perceptions of what is and isn't internet trolling. *Proceedings of the Association for Information Science and Technology*, 55(1), 76–85. <https://doi.org/10.1002/pr2.2018.14505501009>

- Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1217–1230). Portland: CSCW.
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64–73. <https://doi.org/10.2307/3150876>
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588–608. <https://doi.org/10.2307/2094589>
- Coles, B. A., & West, M. (2016a). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior*, 60, 233–244. <https://doi.org/10.1016/j.chb.2016.02.070>
- Coles, B. A., & West, M. (2016b). Weaving the internet together: Imagined communities in newspaper comment threads. *Computers in Human Behavior*, 60, 44–53. <https://doi.org/10.1016/j.chb.2016.02.049>
- Cook, C., Conijn, R., Schaafsma, J., & Antheunis, M. (2019). For whom the gamer trolls: A study of trolling interactions in the online gaming context. *Journal of Computer-Mediated Communication*. <https://doi.org/10.1093/jcmc/zmz014>
- Cook, C., Schaafsma, J., & Antheunis, M. (2017). Under the bridge: An in-depth examination of online trolling in the gaming context. *New Media & Society*, 20(9), 3323–3340. <https://doi.org/10.1177/1461444817748578>
- Costello, M., Hawdon, J., & Ratliff, T. N. (2017). Confronting online extremism: The effect of self-help, collective efficacy, and guardianship on being a target for hate speech. *Social Science Computer Review*, 35(5), 587–605. <https://doi.org/10.1177/0894439316666272>
- Craker, N., & March, E. (2016). The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences*, 102, 79–84.
- Cruz, A. G., Seo, Y., & Rex, M. (2018). Trolling in online communities: A practice-based theoretical perspective. *The Information Society*, 34(1), 15–26. <https://doi.org/10.1080/01972243.2017.1391909>
- Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in Survey Data. *Journal of Experimental Social Psychology*, 66, 4–19.

- de Seta, G. (2013). FCJ-167 Spraying, fishing, looking for trouble: The Chinese Internet and a critical perspective on the concept of trolling. *The Fibreculture Journal*, 22(1), 300–318.
- Demsar, V., Brace-Govan, J., Jack, G., & Sands, S. (2021). The social phenomenon of trolling: Understanding the discourse and social practices of online provocation. *Journal of Marketing Management*, 1–33.
- De', R., Pandey, N., & Pal, A. (2020). Impact of digital surge during covid-19 pandemic: A viewpoint on research and Practice. *International Journal of Information Management*, 55, 102171. <https://doi.org/10.1016/j.ijinfomgt.2020.102171>
- Diamantopoulos, A., & Sigauw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, 17(4), 263–282. <https://doi.org/10.1111/j.1467-8551.2006.00500.x>
- DiFranco, R. (2020). I wrote this paper for the Lulz: The ethics of Internet trolling. *Ethical Theory and Moral Practice*, 23(5), 931–945. <https://doi.org/10.1007/s10677-020-10115-x>
- Dineva, D. P., & Breitsohl, J. C. (2021). Managing trolling in online communities: An organizational perspective. *Internet Research*, 32(1), 292-311.
- Dineva, D. P., Breitsohl, J. C., & Garrod, B. (2017). Corporate conflict management on social media brand fan pages. *Journal of Marketing Management*, 33(9-10), 679–698. <https://doi.org/10.1080/0267257x.2017.1329225>
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, 18(2), 192–203. <https://doi.org/10.1037/1040-3590.18.2.192>
- Dooley, J. J., Pyżalski, J., & Cross, D. (2010). Cyberbullying Versus Face-to-Face Bullying. *Journal of Psychology*, 217(4), 182–188. <https://doi.org/10.1027/0044-3409.217.4.182>
- Duggan, M. (2017, July 11). *Online harassment 2017*. Pew Research Center: Internet, Science & Tech. Retrieved July 14, 2022, from <https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/>
- Dunn, A. M., Heggstad, E. D., Shanock, L. R., & Theilgard, N. (2018). Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences. *Journal of Business and Psychology*, 33(1), 105-121.

- Dynel, M. (2016). "Trolling is not stupid": Internet trolling as the art of deception serving entertainment. *Intercultural Pragmatics*, 13(3). <https://doi.org/10.1515/ip-2016-0015>
- Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 81(6), 358-361.
- Edvardsson, B., Tronvoll, B., & Gruber, T. (2011). Expanding understanding of service exchange and value co-creation: A social construction approach. *Journal of the Academy of Marketing Science*, 39(2), 327–339. <https://doi.org/10.1007/s11747-010-0200-y>
- El-Khoury, J., Haidar, R., Kanj, R. R., Bou Ali, L., & Majari, G. (2021). Characteristics of social media 'detoxification' in university students. *Libyan Journal of Medicine*, 16(1), 1–7. <https://doi.org/10.1080/19932820.2020.1846861>
- Elcessor, E. (2016). Cyborg hoaxes: Disability, deception, and critical studies of digital media. *New Media & Society*, 19(11), 1761–1777. <https://doi.org/10.1177/1461444816642754>
- Ewing, M. T., Wagstaff, P. E., & Powell, I. H. (2013). Brand rivalry and community conflict. *Journal of Business Research*, 66(1), 4–12. <https://doi.org/10.1016/j.jbusres.2011.07.017>
- Festl, R., & Quandt, T. (2013). Social Relations and cyberbullying: The influence of individual and structural attributes on victimization and perpetration via the internet. *Human Communication Research*, 39(1), 101–126. <https://doi.org/10.1111/j.1468-2958.2012.01442.x>
- Festl, R., Scharnow, M., & Quandt, T. (2013). Peer influence, Internet use and cyberbullying: A comparison of different context effects among German adolescents. *Journal of Children and Media*, 7(4), 446–462. <https://doi.org/10.1080/17482798.2013.781514>
- Fichman, P., & Peters, E. (2019). The impacts of territorial communication norms and composition on online trolling. *International Journal of Communication*, 13, 1016–1035.
- Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>

- Frankowski, S., Lupo, A. K., Smith, B. A., Dane'El, M., Ramos, C., & Morera, O. F. (2016). Developing and testing a scale to measure need for drama. *Personality and Individual Differences, 89*, 192–201. <https://doi.org/10.1016/j.paid.2015.10.009>
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research, 25*(2), 186–192. <https://doi.org/10.1177/002224378802500207>
- Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics, 12*(2), 167–185. <https://doi.org/10.1080/19331681.2014.997416>
- Glasø, L., Matthiesen, S. B., Nielsen, M. B., & Einarsen, S. (2007). Do targets of workplace bullying portray a general victim personality profile? *Scandinavian Journal of Psychology, 48*(4), 313–319. <https://doi.org/10.1111/j.1467-9450.2007.00554.x>
- 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.
- Golf-Papez, M., & Veer, E. (2022). Feeding the trolling: Understanding and mitigating online trolling behavior as an unintended consequence. *Journal of Interactive Marketing, 57*(1), 90–114. <https://doi.org/10.1177/10949968221075315>
- Graham, E. (2019). Boundary maintenance and the origins of trolling. *New Media & Society, 21*(9), 2029–2047. <https://doi.org/10.1177/1461444819837561>
- Greszki, R., Meyer, M., & Schoen, H. (2015). Exploring the effects of removing “too fast” responses and respondents from web surveys. *Public Opinion Quarterly, 79*(2), 471–503. <https://doi.org/10.1093/poq/nfu058>
- Griffiths, M. D. (2014). Adolescent trolling in online environments: A brief overview. *Education and Health, 32*(3), 85–87.
- Guan, X., Gong, J., Xie, L., & Huan, T.-C. (2020). Scale development of value co-destruction behavior in tourism. *Tourism Management Perspectives, 36*, 100757. <https://doi.org/10.1016/j.tmp.2020.100757>
- Guo, S. (2016). A meta-analysis of the predictors of cyberbullying perpetration and victimization. *Psychology in the Schools, 53*(4), 432–453. <https://doi.org/10.1002/pits.21914>
- Gupta, B., Iyer, L. S., & Weisskirch, R. S. (2010). Facilitating global E-commerce: A comparison of consumers' willingness to disclose personal information online in the U.S. and in India. *Journal of Electronic Commerce Research, 11*(1), 41–52.

- Guèvremont, A. (2019). Brand hypocrisy from a consumer perspective: Scale Development and validation. *Journal of Product & Brand Management*, 28(5), 598–613. <https://doi.org/10.1108/jpbm-06-2017-1504>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective* (7th ed.). Pearson.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling* (2nd ed.). Sage.
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research. Language, Behaviour, Culture*, 6(2), 215-242.
- Hardaker, C. (2013). Uh. . . not to be nitpicky,,,,,but...the past tense of drag is dragged, not drug. *Journal of Language Aggression and Conflict*, 1(1), 58–86. <https://doi.org/10.1075/jlac.1.1.04har>
- Hassan, M., & Casaló Ariño, L. V. (2016). Consumer devotion to a different height: How consumers are defending the brand within Facebook brand communities. *Internet Research*, 26(4), 963–981. <https://doi.org/10.1108/intr-03-2015-0090>
- Haverila, K. C. (2023). *Trolling Behaviors and Victimization in Online Brand Communities* [Unpublished doctoral dissertation]. Concordia University.
- Haverila, M., McLaughlin, C., Haverila, K. C., & Viskovich, J. (2020). Brand community motives and their impact on brand community engagement: Variations between diverse audiences. *Management Decision*, 59(9), 2286–2308. <https://doi.org/10.1108/md-06-2020-0683>
- Hawdon, J., Oksanen, A., & Räsänen, P. (2014). Victims of online hate groups: American youth's exposure to online hate speech. In J. Hawdon, J. Ryan, & M. Lucht (Eds.), *The causes and consequences of group violence: From bullies to terrorists* (pp. 165-182). Lexington Books.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hong, F.-Y., & Cheng, K.-T. (2018). Correlation between university students' online trolling behavior and online trolling victimization forms, current conditions, and personality

- traits. *Telematics and Informatics*, 35(2), 397–405.
<https://doi.org/10.1016/j.tele.2017.12.016>
- Hopkinson, C. (2013). Trolling in online discussions: From provocation to community-building. *Brno Studies in English*, 39(1), 5–25. <https://doi.org/10.5817/bse2013-1-1>
- Horne, B. D. (2023). Is automated content moderation going to solve our misinformation problems? *SSRN Electronic Journal*, 3(1), 1–5. <https://doi.org/10.2139/ssrn.4359981>
- Howard, K., Zolnierek, K. H., Critz, K., Dailey, S., & Ceballos, N. (2019). An examination of psychosocial factors associated with malicious online trolling behaviors. *Personality and Individual Differences*, 149, 309–314. <https://doi.org/10.1016/j.paid.2019.06.020>
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27(1), 99–114. <https://doi.org/10.1007/s10869-011-9231-8>
- Indiana University. (2021). *Home*. Rob Kling Center for Social Informatics. Retrieved July 2, 2022, from <https://rkcsi.luddy.indiana.edu/>
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1-2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
- Jane, E. A. (2015). Flaming? What flaming? The pitfalls and potentials of researching online hostility. *Ethics and Information Technology*, 17(1), 65-87.
- Japutra, A., Ekinci, Y., & Simkin, L. (2014). Exploring brand attachment, its determinants and outcomes. *Journal of Strategic Marketing*, 22(7), 616–630.
<https://doi.org/10.1080/0965254x.2014.914062>
- Johnson, A. R., Matear, M., & Thomson, M. (2011). 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. <https://doi.org/10.1086/657924>
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, 22(2), 420–432. <https://doi.org/10.1037/a0019265>
- Jung, N. Y., Kim, S., & Kim, S. (2014). Influence of consumer attitude toward online brand community on Revisit Intention and Brand Trust. *Journal of Retailing and Consumer Services*, 21(4), 581–589. <https://doi.org/10.1016/j.jretconser.2014.04.002>
- Jussinoja, T. (2018). *Lifecycle Of Internet Trolls* [Doctoral dissertation, University of Jyväskylä]. JYX Digital Repository.

- Kaakinen, M., Keipi, T., Oksanen, A., & Räsänen, P. (2018). How does social capital associate with being a victim of online hate? Survey evidence from the United States, the United Kingdom, Germany, and Finland. *Policy & Internet, 10*(3), 302–323. <https://doi.org/10.1002/poi3.173>
- Kenny, D. A. (2016, April 9). *Multiple Latent Variable Models: Confirmatory Factor Analysis*. Homepage of David A. Kenny. Retrieved April 20, 2023, from <http://davidakenny.net/cm/mfactor.htm>
- Kim, K., Byon, K. K., & Baek, W. (2020). Customer-to-customer value co-creation and co-destruction in sporting events. *The Service Industries Journal, 40*(9-10), 633–655. <https://doi.org/10.1080/02642069.2019.1586887>
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.
- Kock, N. (2015). Common method bias in PLS-SEM. *International Journal of e-Collaboration, 11*(4), 1–10. <https://doi.org/10.4018/ijec.2015100101>
- Labrecque, L. I., Markos, E., Yuksel, M., & Khan, T. A. (2022). Value creation (vs value destruction) as an unintended consequence of negative comments on [innocuous] brand social media posts. *Journal of Interactive Marketing, 57*(1), 115–140. <https://doi.org/10.1177/10949968221075820>
- Ladanyi, J., & Doyle-Portillo, S. (2017). The development and validation of the grief play scale (GPS) in mmorpgs. *Personality and Individual Differences, 114*, 125–133. <https://doi.org/10.1016/j.paid.2017.03.062>
- Lamb, R., & Kling, R. (2003). Reconceptualizing users as social actors in Information Systems Research. *MIS Quarterly, 27*(2), 197–235. <https://doi.org/10.2307/30036529>
- Leaver, T. (2013). FCJ-163 Olympic trolls: Mainstream memes and digital discord? *The Fibreculture Journal, 22*(1), 216–233.
- Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology, 86*(1), 114–121. <https://doi.org/10.1037/0021-9010.86.1.114>
- Lisjak, M., Lee, A. Y., & Gardner, W. L. (2012). When a threat to the brand is a threat to the self. *Personality and Social Psychology Bulletin, 38*(9), 1120–1132. <https://doi.org/10.1177/0146167212445300>
- Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *An Introduction to Statistical Concepts* (3rd ed.). Routledge.

- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety & Stress Scales* (2nd ed.). Psychology Foundation of Australia.
- MacInnis, D. (2017). Developing Conceptual Articles for JCR (Winter 2016/2017). *Journal of Consumer Research*. Retrieved July 2, 2022, from https://academic.oup.com/DocumentLibrary/JCR/Developing_Conceptual_Articles_Research_Curations_JCR_MacInnis_Final.pdf
- Maltby, J., Day, L., Hatcher, R. M., Tazzyman, S., Flowe, H. D., Palmer, E. J., ... Cutts, K. (2015). Implicit theories of online trolling: Evidence that attention-seeking conceptions are associated with increased psychological resilience. *British Journal of Psychology*, *107*(3), 448–466. <https://doi.org/10.1111/bjop.12154>
- March, E., Grieve, R., Marrington, J., & Jonason, P. K. (2017). Trolling on tinder® (and other dating apps): Examining the role of the dark tetrad and impulsivity. *Personality and Individual Differences*, *110*, 139–143. <https://doi.org/10.1016/j.paid.2017.01.025>
- March, E., & Marrington, J. (2019). A qualitative analysis of internet trolling. *Cyberpsychology, Behavior, and Social Networking*, *22*(3), 192-197.
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data&Society. Retrieved March 1, 2023, from https://www.datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf
- Masui, K. (2019). Loneliness moderates the relationship between dark tetrad personality traits and internet trolling. *Personality and Individual Differences*, *150*, 109475. <https://doi.org/10.1016/j.paid.2019.06.018>
- Matias, J. N. (2019). The Civic Labor of volunteer moderators online. *Social Media + Society*, *5*(2), 1–12. <https://doi.org/10.1177/2056305119836778>
- McAloon, J. (2015, September 30). *Why do celebrities like Lena Dunham leave Twitter?* The Telegraph. Retrieved March 2, 2023, from <https://www.telegraph.co.uk/culture/tvandradio/11537494/Why-do-celebrities-leave-Twitter.html>
- McClain, C., Vogels, E. A., Perrin, A., Sechopoulos, S., & Rainie, L. (2021, September 1). *The internet and the Pandemic*. Pew Research Center: Internet, Science & Tech. <https://www.pewresearch.org/internet/2021/09/01/the-internet-and-the-pandemic/>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437-455.

- Miller, V. (2012). A Crisis of presence: On-line culture and being in the world. *Space and Polity*, 16(3), 265–285. <https://doi.org/10.1080/13562576.2012.733568>
- Nepomuceno, M. V., Rahemi, H., Cenesizoglu, T., & Charlin, L. (2023). Should we feed the trolls? using marketer-generated content to explain average toxicity and product usage. *Journal of Interactive Marketing*. <https://doi.org/10.1177/10949968231172153>
- Nepomuceno, M. V., Visconti, L. M., & Cenesizoglu, T. (2020). A model for investigating the impact of owned social media content on commercial performance and its application in large and mid-sized online communities. *Journal of Marketing Management*, 36(17-18), 1762–1804. <https://doi.org/10.1080/0267257x.2020.1825112>
- November, P. (2004). Seven reasons why marketing practitioners should ignore marketing academic research. *Australasian Marketing Journal*, 12(2), 39–50. [https://doi.org/10.1016/s1441-3582\(04\)70096-8](https://doi.org/10.1016/s1441-3582(04)70096-8)
- Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, 39(9), 745-750.
- Olson, K. (2010). An examination of questionnaire evaluation by expert reviewers. *Field Methods*, 22(4), 295–318. <https://doi.org/10.1177/1525822x10379795>
- Olweus, D. (2014). Victimization by peers: Antecedents and long-term outcomes. In K. H. Rubin, & J. B. Asendorpf (Eds.), *Social Withdrawal, Inhibition, and Shyness* (pp. 325-352). Psychology Press.
- Ordabayeva, N., Cavanaugh, L. A., & Dahl, D. W. (2022). The upside of negative: Social distance in online reviews of identity-relevant brands. *Journal of Marketing*, 86(6), 70–92. <https://doi.org/10.1177/00222429221074704>
- Ounvorawong, N. (2021). *Brand victimisation: When consumers are bullied by fellow brand followers in online brand communities* [Doctoral dissertation, University of Kent]. Kent Academic Repository.
- Ounvorawong, N., Breitsohl, J. C., Lowe, B., & Laffey, D. (2022). Outcomes of cyber-victimization and bystander reactions in online brand communities. *International Journal of Electronic Commerce*, 26(2), 200–221. <https://doi.org/10.1080/10864415.2022.2050582>
- Paavola, J., Helo, T., Jalonen, H., Sartonen, M., & Huhtinen, A.-M. (2016). Understanding the trolling phenomenon: The automated detection of bots and cyborgs in the social media. *Journal of Information Warfare*, 15(4), 100–111.

- Park, S., Na, E.-Y., & Kim, E.-M. (2014). The relationship between online activities, netiquette and cyberbullying. *Children and Youth Services Review, 42*, 74–81. <https://doi.org/10.1016/j.childyouth.2014.04.002>
- Patchin, J. W., & Hinduja, S. (2010). Cyberbullying and self-esteem. *Journal of School Health, 80*(12), 614–621. <https://doi.org/10.1111/j.1746-1561.2010.00548.x>
- Patchin, J. W., & Hinduja, S. (2011). Traditional and nontraditional bullying among youth: A test of general strain theory. *Youth & Society, 43*(2), 727–751. <https://doi.org/10.1177/0044118x10366951>
- Patchin, J. W., & Hinduja, S. (2015). Measuring cyberbullying: Implications for research. *Aggression and Violent Behavior, 23*, 69–74. <https://doi.org/10.1016/j.avb.2015.05.013>
- Paulhus, D. L., Buckels, E. E., Trapnell, P. D., & Jones, D. N. (2021). Screening for dark personalities. *European Journal of Psychological Assessment, 37*(3), 208–222. <https://doi.org/10.1027/1015-5759/a000602>
- Pavlou, P. A., Liang, H., & Xue, Y. (2007). Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective. *MIS Quarterly, 31*(1), 105–136. <https://doi.org/10.2307/25148783>
- Pelley, S. (2021, October 4). *Whistleblower: Facebook is misleading the public on progress against hate speech, violence, misinformation*. CBS News. <https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/>
- Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing, 19*(1), 27–41. <https://doi.org/10.1509/jppm.19.1.27.16941>
- Plé, L., & Cáceres, R. C. (2010). Not always co-creation: Introducing interactional co-destruction of value in service-dominant logic. *Journal of Services Marketing, 24*(6), 430–437. <https://doi.org/10.1108/08876041011072546>
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>
- 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), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>

- Pons, F., Mourali, M., & Nyeck, S. (2006). Consumer orientation toward sporting events: Scale development and validation. *Journal of Service Research*, 8(3), 276–287. <https://doi.org/10.1177/1094670505283931>
- Prahalad, C. K., & Ramaswamy, V. (2004). *The Future of Competition: Co-creating Unique Value with Customers*. Harvard Business School Press.
- Relling, M., Schnittka, O., Sattler, H., & Johnen, M. (2016). Each can help or hurt: Negative and positive word of mouth in social network brand communities. *International Journal of Research in Marketing*, 33(1), 42–58. <https://doi.org/10.1016/j.ijresmar.2015.11.001>
- Rotman, J. D., Khamitov, M., & Connors, S. (2017). Lie, cheat, and steal: How harmful brands motivate consumers to act unethically. *Journal of Consumer Psychology*, 28(2), 353–361. <https://doi.org/10.1002/jcpy.1002>
- Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior*, 58, 461–470. <https://doi.org/10.1016/j.chb.2016.01.022>
- Sanfilippo, M. R., Fichman, P., & Yang, S. (2018). Multidimensionality of online trolling behaviors. *The Information Society*, 34(1), 27-39.
- Savoldi, F., & Ferraz de Abreu, P. (2016). Bullying, cyberbullying and Internet usage among young people in post-conflict Belfast. *Cogent Social Sciences*, 2(1), 1132985. <https://doi.org/10.1080/23311886.2015.1132985>
- Sawyer, S., & Tyworth, M. (2006) Social Informatics: An Information Society for All? In T. Berleur, M.I. Numinen & J. Impagliazzo (Eds.), *IFIP International Federation for Information Processing*. (pp. 49-62). Springer.
- Shachaf, P., & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science*, 36(3), 357–370. <https://doi.org/10.1177/0165551510365390>
- Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation. *Organizational Research Methods*, 19(1), 80–110. <https://doi.org/10.1177/1094428115598239>
- Shin, H., & Larson, L. R. L. (2020). The bright and dark sides of humorous response to online customer complaint. *European Journal of Marketing*, 54(8), 2013–2047. <https://doi.org/10.1108/ejm-08-2018-0522>
- Steffgen, G., König, A., Pfetsch, J., & Melzer, A. (2011). Are cyberbullies Less empathic? Adolescents' cyberbullying behavior and empathic responsiveness. *Cyberpsychology*,

- Behavior, and Social Networking*, 14(11), 643–648.
<https://doi.org/10.1089/cyber.2010.0445>
- Stewart, R. W., Drescher, C. F., Maack, D. J., Ebesutani, C., & Young, J. (2014). The development and psychometric investigation of the cyberbullying scale. *Journal of Interpersonal Violence*, 29(12), 2218–2238.
<https://doi.org/10.1177/0886260513517552>
- Stieler, M., Weismann, F., & Germelmann, C. C. (2014). Co-destruction of value by spectators: The case of silent protests. *European Sport Management Quarterly*, 14(1), 72–86. <https://doi.org/10.1080/16184742.2013.865249>
- Sun, L. H., & Fichman, P. (2019). The collective trolling lifecycle. *Journal of the Association for Information Science and Technology*, 71(7), 770–783.
- Synnott, J., Coulias, A., & Ioannou, M. (2017). Online trolling: The case of Madeleine McCann. *Computers in Human Behavior*, 71, 70–78.
<https://doi.org/10.1016/j.chb.2017.01.053>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson.
- Thacker, S., & Griffiths, M. D. (2012). An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning*, 2(4), 17–33. <https://doi.org/10.4018/ijcbpl.2012100102>
- Thomas, H. J., Scott, J. G., Coates, J. M., & Connor, J. P. (2019). Development and validation of the bullying and cyberbullying scale for adolescents: A multi-dimensional measurement model. *British Journal of Educational Psychology*, 89(1), 75–94.
<https://doi.org/10.1111/bjep.12223>
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26(3), 277–287. <https://doi.org/10.1016/j.chb.2009.11.014>
- United Nations. (2023). *The impact of Digital Technologies*. United Nations.
<https://www.un.org/en/un75/impact-digital-technologies>
- van Geel, M., Goemans, A., Toprak, F., & Vedder, P. (2017). Which personality traits are related to traditional bullying and cyberbullying? A study with the Big Five, Dark Triad and sadism. *Personality and Individual Differences*, 106, 231–235.
<https://doi.org/10.1016/j.paid.2016.10.063>
- van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods*, 12(2), 368–392.
<https://doi.org/10.1177/1094428107309322>

- Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, 68(1), 1–17. <https://doi.org/10.1509/jmkg.68.1.1.24036>
- Vargo, S. L., & Lusch, R. F. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, 44(1), 5–23. <https://doi.org/10.1007/s11747-015-0456-3>
- Wilson, N. C., & Seigfried-Spellar, K. C. (2022). Cybervictimization, social, and financial strains influence internet trolling behaviors: A general strain theory perspective. *Social Science Computer Review*, 1–16. <https://doi.org/10.1177/08944393211065868>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research. *The Counseling Psychologist*, 34(6), 806-838.
- Wu, B., Li, F., Zhou, L., Liu, M., & Geng, F. (2021). Are mindful people less involved in online trolling? A moderated mediation model of perceived social media fatigue and moral disengagement. *Aggressive Behavior*, 48(3), 309–318. <https://doi.org/10.1002/ab.22013>
- Yi, Y., & Gong, T. (2013). Customer value co-creation behavior: Scale development and validation. *Journal of Business Research*, 66(9), 1279–1284. <https://doi.org/10.1016/j.jbusres.2012.02.026>
- YouGov. (2014, October). *YouGov survey results*. YouGov. Retrieved March 1, 2023, from http://cdn.yougov.com/cumulus_uploads/document/3yqcx5p1jb/InternalResults_141021_posting_online_and_trolls_Website.pdf
- Zahrai, K., Veer, E., Ballantine, P. W., & de Vries, H. P. (2022). A conceptual framework for managing excessive social media use. In K. Kubacki, L. Parker, C. Domegan, & L. Brennan (Eds.), *The Routledge Companion to Marketing and Society* (pp. 156-175). Routledge.
- Zhang, D., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2), 456–481. <https://doi.org/10.1080/07421222.2016.1205907>
- Öberseder, M., Schlegelmilch, B. B., Murphy, P. E., & Gruber, V. (2014). Consumers' perceptions of corporate social responsibility: Scale development and validation. *Journal of Business Ethics*, 124(1), 101–115. <https://doi.org/10.1007/s10551-013-1787-y>

Appendices

Appendix A: Summary of Research Propositions

Research Propositions	Level of Analysis	Focus of Analysis
<i>RP1: When shifting from the general trolling construct to the brand trolling construct, the level of analysis has expanded from a triadic perspective to a multi-level perspective that includes (a) micro-, (b) meso- and (c) macro-level considerations.</i>	Individual, Community, Organization/Brand	Level of Analysis
<i>RP2: (a) Sadism, (b) psychopathy, and (c) depression positively predict brand trolling.</i>	Individual	Predictors
<i>RP3: Agreeableness negatively predicts brand trolling.</i>	Individual	Predictors
<i>RP4: (a) High social capital, (b) social introversion, (c) depression, and (d) a willingness to divulge personal information positively predict brand trolling victimization.</i>	Individual	Predictors
<i>RP5: (a) The immediacy of brand community interactions, (b) the visibility of brand community posts, (c) customers' fast access to brand-related information, and (d) a brand community's closeness positively predict brand trolling victimization.</i>	Community	Predictors
<i>RP6: (a) A brand's perceived harmfulness, (b) brand hypocrisy, (c) brand hate, and (d) moral violation positively predict brand trolling victimization.</i>	Organization/Brand	Predictors
<i>RP7: (a) Brand trustworthiness, and (b) brand authenticity negatively predict brand trolling victimization.</i>	Organization/Brand	Predictors
<i>RP8: Brand trolling victimization positively predicts (a) feelings of hostility, and (b) brand trolling behavior.</i>	Individual	Outcomes
<i>RP9: Brand trolling victimization negatively affects the (a) efficiency and (b) well-being of brand community employees.</i>	Individual	Outcomes
<i>RP10: Brand trolling victimization negatively predicts community-related outcomes such as community members' (a) joining behaviors, (b) level of engagement, (c) quality of engagement, (d) brand community satisfaction levels, (e) visiting and (f) revisiting intentions, (g) community trust levels, and (h) positive WOM intentions towards the brand community.</i>	Community	Outcomes

Research Propositions	Level of Analysis	Focus of Analysis
<i>RP11: Brand trolling victimization negatively predicts brand-related outcomes such as (a) customer brand attitudes, (b) brand reputation, (c) positive WOM intentions towards the brand, (d) brand image, (e) brand trust, (f) brand satisfaction, and (g) brand purchase intentions.</i>	Organization/Brand	Outcomes
<i>RP12: Brand trolling contributes to value co-destruction on the individual-, community-, and organization-level.</i>	Individual, Community, Organization/Brand	Outcomes
<i>RP13: Brand trolling contributes to value co-creation on the individual-, community-, and organization-level.</i>	Individual, Community, Organization/Brand	Outcomes
<i>RP14: When shifting from the general trolling construct to the brand trolling construct, the roles and expectations have significantly changed for (a) the targets of trolling, (b) the community members, (c) the trolls themselves, and (d) the employees.</i>	Individual	Roles
<i>RP15: Brand trolling differs in its frequency, severity, and manifestation across online communities with (a) different norms, rules, and expectations, (b) different levels of experience with the behavior, (c) different goals, (d) varying levels of “officialness”, (e) closeness, and (f) community size.</i>	Community	Context
<i>RP16: Brand trolling differs in its frequency, severity, and manifestation across online communities of brands (a) with varying user base sizes, and with varying levels of (b) involvement and (c) self-relevance.</i>	Organization/Brand	Context

Appendix B: Summary of Research Questions

Research Questions	Level of Analysis	Focus of Analysis
<i>RQ1: What specific value co-creation outcomes emerge from brand trolling in online brand communities?</i>	Individual, Community, Organization/Brand	Outcomes
<i>RQ2: What specific value co-destruction outcomes emerge from brand trolling in online brand communities?</i>	Individual, Community, Organization/Brand	Outcomes
<i>RQ3: What constructs could potentially act as moderators and/or mediators in a brand trolling framework?</i>	Individual, Community, Organization/Brand	Mediators and Moderators
<i>RQ4: When developing a brand trolling behavior scale, what items and dimensions should be included in the scale to reflect what brand trolling accurately is?</i>	-	Scale Development
<i>RQ5: What are the positive consequences that can result from brand trolling and/or brand trolling management?</i>	Individual, Community, Organization/Brand	Outcomes
<i>RQ6: What are the relevant predictors, outcomes and contextual elements of brand trolling on the meta-level (i.e., the ecosystem level)?</i>	Ecosystem	Predictors, Outcomes, Context
<i>RQ7: What are the strategies and goals of troll marketing?</i>	Organization/Brand	Troll Marketing
<i>RQ8: When directed towards a brand or an organization instead of an individual, how does trolling differ in its relationships, contextual elements, frequency, severity and manifestation?</i>	Individual, Community, Organization/Brand	Individual-Directed vs. Brand-Directed Trolling
<i>RQ9: How would the public's perceptions differ when viewing brand-directed trolling (where the targets include organizations and brands) compared to individual-directed trolling?</i>	Individual, Community, Organization/Brand	Individual-Directed vs. Brand-Directed Trolling, Perceptions
<i>RQ10: What is the relationship between trolling management strategies and community-related factors such as community members' (a) joining behaviors (b) level of engagement, (c) quality of engagement, (d) brand community satisfaction levels, (e) visiting and (f) revisiting intentions, (g) community trust levels, and (h) positive WOM intentions towards the brand community?</i>	Community	Management of Behavior: Outcomes

Research Questions	Level of Analysis	Focus of Analysis
<i>RQ11: What is the relationship between trolling management strategies and brand-related factors such as (a) customer brand attitudes, (b) brand reputation, (c) positive WOM intentions towards the brand, (d) brand image, (e) brand trust, (f) brand satisfaction, and (g) brand purchase intentions?</i>	Organization /Brand	Management of Behavior: Outcomes

Appendix C: Distraction Items in Trolling Behavior Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have deliberately...</i>					
1. Caused distractions online					
2. Caused inconveniences for individuals online					
3. Caused disruptions online					
4. Caused confusion online					
5. Caused disturbances online					
6. Caused dysfunction online					
7. Created disorder online					
8. Created mischief online					
9. Wasted an individual's time online					
10. Sent online messages repeatedly to annoy individuals					
11. Sent irrelevant messages to online discussions					
12. Derailed online discussions					
13. Lured individuals into pointless arguments online					

Note: The items in bold were retained in the final 15-item trolling behavior scale

Appendix D: Deception Items in Trolling Behavior Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have deliberately...</i>					
1. Impersonated individuals online					
2. Misrepresented myself online					
3. Provided false information online					
4. Deceived individuals online					
5. Misled individuals online					
6. Lied to individuals online					
7. Manipulated individuals online					
8. Created false online profiles					
9. Created fake online reviews					

Note: The items in bold were retained in the final 15-item trolling behavior scale

Appendix E: Provocation Items in Trolling Behavior Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have deliberately...</i>					
1. Upset individuals online					
2. Annoyed individuals online					
3. Angered individuals online					
4. Bothered individuals online					
5. Antagonized individuals online					
6. Been inflammatory to individuals online					
7. Taunted individuals online					
8. Trolled individuals online					
9. Mocked individuals online					
10. Made individuals react negatively online					
11. Provoked individuals online					
12. Made individuals respond negatively to me online					
13. Been overly argumentative with individuals online					
14. Been controversial in online communication					
15. Caused arguments online					
16. Escalated arguments online					
17. Created conflict online					
18. Caused individuals to fight online					
19. Caused outrage online					
20. Stirred up drama online					

Note: The items in bold were retained in the final 15-item trolling behavior scale

Appendix F: Distraction Items in Trolling Victimization Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have experienced individuals deliberately...</i>					
1. Causing distractions online					
2. Causing inconveniences for others/me online					
3. Causing disruptions online					
4. Causing confusion online					
5. Causing disturbances online					
6. Causing dysfunction online					
7. Creating disorder online					
8. Creating mischief online					
9. Wasting others'/my time online					
10. Sending repeated online messages to annoy others/me					
11. Sending irrelevant messages to online discussions					
12. Derailing online discussions					
13. Luring others/me into pointless arguments online					

Note: The items in bold were retained in the final 16-item trolling victimization scale

Appendix G: Deception Items in Trolling Victimization Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have experienced individuals deliberately...</i>					
1. Impersonating others/me online					
2. Misrepresenting themselves online					
3. Providing false information online					
4. Deceiving others/me online					
5. Misleading others/me online					
6. Lying to others/me online					
7. Manipulating others/me online					
8. Creating false online profiles					
9. Creating fake online reviews					

Note: The items in bold were retained in the final 16-item trolling victimization scale

Appendix H: Provocation Items in Trolling Victimization Scale

	Never	Rarely	Sometimes	Often	Always
<i>I have experienced individuals deliberately...</i>					
1. Upsetting others/me online					
2. Annoying others/me online					
3. Angering others/me online					
4. Bothering others/me online					
5. Antagonizing others/me online					
6. Being inflammatory to others/me online					
7. Taunting others/me online					
8. Trolling others/me online					
9. Mocking others/me online					
10. Making others/me react negatively online					
11. Provoking others/me online					
12. Wanting negative responses online					
13. Being overly argumentative with others/me online					
14. Been controversial in online communication					
15. Causing arguments online					
16. Escalating arguments online					
17. Creating conflict online					
18. Causing others/me to fight online					
19. Causing outrage online					
20. Stirring up drama online					

Note: The items in bold were retained in the final 16-item trolling victimization scale

Appendix I: List of Scales to Test Predictive Validity for Trolling Behavior

Construct	Items
<p>Sadism (Paulhus et al., 2021)</p>	<p><i>Please indicate your level of agreement or disagreement with the following statements:</i></p> <ol style="list-style-type: none"> 1. Watching a fist-fight excites me 2. I really enjoy violent films and video games 3. It's funny when idiots fall flat on their face 4. I enjoy watching violent sports 5. Some people deserve to suffer 6. Just for kicks, I've said mean things on social media 7. I know how to hurt someone with words alone
<p>Depression (Lovibond & Lovibond, 1995)</p>	<p><i>Please read each statement and indicate how much the statement applied to you in the past month:</i></p> <ol style="list-style-type: none"> 1. I couldn't seem to experience any positive feelings at all 2. I found it difficult to work up the initiative to do things 3. I felt that I had nothing to look forward to 4. I felt down-hearted and blue 5. I was unable to become enthusiastic about anything 6. I felt I wasn't worth much as a person 7. I felt that life was meaningless
<p>Psychopathy (Jonason & Webster, 2010)</p>	<p><i>I tend to...</i></p> <ol style="list-style-type: none"> 1. Lack remorse 2. Be unconcerned with the morality of my actions 3. Be callous or insensitive 4. Be cynical
<p>Personality (Donnellan et al., 2006) E = Extraversion A = Agreeableness C = Conscientiousness N = Neuroticism I = Intellect/Imagination R = Reverse Scored Item</p>	<p><i>Please indicate your level of agreement or disagreement with the following statements:</i></p> <ol style="list-style-type: none"> 1. I am the life of the party (E) 2. I sympathize with others' feelings (A) 3. I get chores done right away (C) 4. I have frequent mood swings (N) 5. I have a vivid imagination (I) 6. I don't talk a lot (E) (R) 7. I am not interested in other people's problems (A) (R) 8. I often forget to put things back in their proper place (C) (R) 9. I am relaxed most of the time (N) (R) 10. I am not interested in abstract ideas (I) (R) 11. I talk to a lot of different people at parties (E) 12. I feel others' emotions (A) 13. I like order (C) 14. I get upset easily (N)

Construct	Items
	15. I have difficulty understanding abstract ideas (I) (R) 16. I keep in the background (E) (R) 17. I am not really interested in others (A) (R) 18. I make a mess of things (C) (R) 19. I seldom feel blue (N) (R) 20. I do not have a good imagination (I) (R)
Cyberbullying (Patchin & Hinduja, 2015)	<p><i>Please indicate how often you have engaged in each of the following statements:</i></p> <ol style="list-style-type: none"> 1. I cyberbullied others 2. I posted mean or hurtful comments about someone online 3. I posted a mean or hurtful picture online of someone 4. I posted a mean or hurtful video online of someone 5. I spread rumors about someone online 6. I threatened to hurt someone online 7. I threatened to hurt someone through a cell phone text message 8. I created a mean or hurtful web page about someone 9. I pretended to be someone else online and acted in a way that was mean or hurtful to them
GAIT (Buckels et al., 2014)	<p><i>Please indicate how often you have engaged in each of the following statements:</i></p> <ol style="list-style-type: none"> 1. I have sent people to shock websites for the lulz 2. I like to troll people in forums or the comments section of websites 3. I enjoy griefing other plays in multiplayer games 4. The more beautiful and pure a thing is, the more satisfying it is to corrupt

Appendix J: List of Scales to Test Predictive Validity for Trolling Victimization

Construct	Items
<p>Feelings of Hostility (Anderson & Carnagey, 2009)</p>	<p><i>After the trolling incident, I felt...</i></p> <ol style="list-style-type: none"> 1. Offended 2. Angry 3. Outraged 4. Enraged 5. Mad 6. Aggravated 7. Frustrated 8. Irritable 9. Furious
<p>Depression (Lovibond & Lovibond, 1995)</p>	<p><i>Please read each statement and indicate how much the statement applied to you in the past month:</i></p> <ol style="list-style-type: none"> 1. I couldn't seem to experience any positive feelings at all 2. I found it difficult to work up the initiative to do things 3. I felt that I had nothing to look forward to 4. I felt down-hearted and blue 5. I was unable to become enthusiastic about anything 6. I felt I wasn't worth much as a person 7. I felt that life was meaningless
<p>Willingness to Divulge Personal Information (Gupta et al., 2010; Phelps et al., 2000)</p>	<p><i>Please indicate your willingness to provide the following pieces of information:</i></p> <ol style="list-style-type: none"> 1. Demographic information (e.g., age, marital status, occupation) 2. Lifestyle information (e.g., hobbies, leisure activities) 3. Media habits (e.g., most visited social media, websites) 4. Financial information (e.g., annual household income) 5. Contact information (e.g., email, phone number, addresses) 6. Names (e.g., first name, family name)

Construct	Items
<p>Personality (Donnellan et al., 2006) E = Extraversion A = Agreeableness C = Conscientiousness N = Neuroticism I = Intellect/Imagination</p> <p>R = Reverse Scored Item</p>	<p><i>Please indicate your level of agreement or disagreement with the following statements:</i></p> <ol style="list-style-type: none"> 1. I am the life of the party (E) 2. I sympathize with others' feelings (A) 3. I get chores done right away (C) 4. I have frequent mood swings (N) 5. I have a vivid imagination (I) 6. I don't talk a lot (E) (R) 7. I am not interested in other people's problems (A) (R) 8. I often forget to put things back in their proper place (C) (R) 9. I am relaxed most of the time (N) (R) 10. I am not interested in abstract ideas (I) (R) 11. I talk to a lot of different people at parties (E) 12. I feel others' emotions (A) 13. I like order (C) 14. I get upset easily (N) 15. I have difficulty understanding abstract ideas (I) (R) 16. I keep in the background (E) (R) 17. I am not really interested in others (A) (R) 18. I make a mess of things (C) (R) 19. I seldom feel blue (N) (R) 20. I do not have a good imagination (I) (R)
<p>Cyberbullying Victimization (Patchin & Hinduja, 2015)</p>	<p><i>Please indicate how often you have experienced each of the following statements:</i></p> <ol style="list-style-type: none"> 1. I have been cyberbullied 2. Someone posted mean or hurtful comments about me online 3. Someone posted a mean or hurtful picture online of me online 4. Someone posted a mean or hurtful video online of me online 5. Someone created a mean or hurtful web page about me 6. Someone spread rumors about me online 7. Someone threatened to hurt me through a cell phone text message 8. Someone threatened to hurt me online 9. Someone pretended to be me online and acted in a way that was mean or hurtful
<p>GAIT Victimization (Buckels et al., 2014)</p>	<p><i>Please indicate how often you have experienced each of the following statements:</i></p> <ol style="list-style-type: none"> 1. I have been sent to shock websites by someone for the lulz 2. I have been trolled by individuals in forums or the comments section of websites 3. I have been grieved by other players in multiplayer games

Construct	Items
	4. I have witnessed online individuals who find satisfaction in corrupting beautiful and pure things