

LET'S AGREE TO AGREE: EFFECTS OF SELF-AWARENESS AND SOCIAL
IDENTITY ON ONLINE DELIBERATION

A Dissertation
presented to
the Faculty of the Graduate School
at the University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

By
MICHELLE E. FUNK
Dr. Elizabeth Behm-Morawitz, Dissertation Supervisor

JULY 2018

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

LET'S AGREE TO AGREE: EFFECTS OF SELF-AWARENESS AND SOCIAL
IDENTITY ON ONLINE DELIBERATION

presented by Michelle Funk,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Elizabeth Behm-Morawitz

Professor Benjamin Warner

Professor J. Brian Houston

Professor Laura Scherer

Acknowledgements

This project started because of my own tendency to get really passionate (to put it flatteringly) in online political fights, and subsequently exhausted. While some people told me “chill, it’s not worth it,” there are some people who never told me that, and it’s these people I thank especially. I think these are the people who see the Internet the way I see it, or as I want to: for its potential—even if it can be taxing to continue leaping into the fray when it seems that, over and over, no footing is gained in the name of... whatever it is you happen to be arguing for.

I want to extend special thanks to the following individuals: Dr. Lissa Behm-Morawitz for being the most incredible advisor, and for making this process fairly painless and actually very exciting. I owe it to Lissa that I never burned out on this project, and continued to look on the bright side the entire way through. I’m not sure how many people can say they completed a dissertation with precisely zero tears, but it’s due to her that I am among them, and that I don’t dread the prospect of expanding upon this work.

I’d also like to thank my committee members: Dr. Benjamin Warner for letting me drop by on short notice with my dry erase board and a million questions, and for loaning me some of the books that helped me reach the majority of the “aha!” moments I experienced in this process. To Dr. Brian Houston for being an example of balancing one’s pure commitment to research, having a million other huge responsibilities, and somehow also being incredibly calm all the time (and for giving me a B on the first, roughest iteration of this project I submitted in his political communication course, which pushed me to work to make it something I’m proud of). And to Dr. Laura Scherer, whose

one-on-one attention helped me realize what my specific theoretical goals are as a researcher, and how to bravely dive into an entirely new body of literature without feeling unwelcome.

To Ethan Morrow, my research assistant (and new recruit to the Master's program in the Department of Communication!) who I literally could not have completed this project without. Ethan provided the lab hours that I could not, so that I could finish data collection with relatively little infringement upon my teaching and other duties. Saint Ethan, truly.

To Dr. Julius Riles, for trusting in me during the most distracted time in my life (dissertation, marriage, job market) to help him in his own research during a high-stakes time in his own career. I got a lot out of being his research assistant—but especially a revitalized appreciation for doing things the right way, not the easy way. Semi-relatedly, I also make a note to emulate certain writers' styles, and he is one of my favorites.

To Martha Crump and Cathy Illingworth, who keep this ship sailing very smoothly, and we hardly ever recognize it because there are never any bumps in the journey.

I would like to extend probably the most thanks to my officemate and colleague Dr. Calvin Coker, for the long lunches and likely hundreds of hours of waxing philosophical. I think our conversations helped me bend my mind around anything I previously thought unattainably difficult, or at the very least you helped me develop the confidence that I *could* tackle anything by insisting that you believed I could.

To my roommate and colleague, Dr. Deirdre Zerilli, and my ~~boyfriend~~ husband (!!!) Stephen. You two see me at my ugliest and you support me anyway. Also a great

thanks to my parents, all of the students I taught while I was writing this, and all of my other friends who don't study what I study, who were probably annoyed with my obsessive inability to talk about anything else over the last few years. Finally, thank you anyone I ever fought with on the Internet, personal acquaintance or entirely anonymous stranger. Those interactions stuck with me, and I'm ever more pleasantly curious about these behaviors because of them. This project isn't even close to finished, because I have so many more questions now—a lifetime of research ahead of me. Believe it, I'm coming back to every single one of you with more questions soon.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	vi
CHAPTER 1 - INTRODUCTION	1
CHAPTER 2 - LITERATURE REVIEW	8
Political Deliberation as a Necessarily Civil Practice	8
Deliberation in the Online Environment	14
Online spaces as a proponent of confirmation bias	15
Demonstrating the self, deliberating with the other	19
Self-Disclosure (and Inviting that from Others) as an Antidote to Online Toxicity	25
Online disinhibition	29
Online disinhibition, self-disclosure, and attentive listening	37
Online disinhibition and outgroup contact	39
Deindividuation and Hostility	43
CHAPTER 3 - METHOD	53
Sample collection	53
Procedure	54
Measures - Independent Variables and Covariates	58
Measures - Dependent	62
Pilot Study	66
Description of Main Study Sample	78
CHAPTER 4 - RESULTS	81
Research Question 1	82
Hypothesis 1 and Research Question 2	85
Hypothesis 2	86
Research Question 3	86
Hypotheses 3 and 4	88
Hypothesis 5 and Research Question 4	90
Research Question 5	93
Research Question 6	94
CHAPTER 5 - DISCUSSION	95
Theoretical and Practical Implications	109
Limitations and Directions for Future Research	112
Concluding Remarks	114
REFERENCES	118
VITA	139

LET'S AGREE TO AGREE: EFFECTS OF SELF-AWARENESS AND
SOCIAL IDENTITY ON ONLINE DELIBERATION

Michelle E. Funk

Dr. Elizabeth Behm-Morawitz, Dissertation Supervisor

ABSTRACT

This experiment assesses the effects of ingroup versus outgroup communication in an online, one-on-one, anonymous setting on perceptions of deliberation quality in a conversation task pertaining to abortion policy. Additionally, the study examines the effects of different secondary goals during conversation tasks—those of persuasion, and those of other-awareness, or understanding—in concert with trait online disinhibition tendencies in individuals. Finally, outcomes related to perceptions of one's conversation partner (civility, intelligence, rationality, hostility, and others), perceived compromise, feelings of political deliberation cynicism, and accuracy of other-awareness (i.e., ability to correctly recall a partner's abortion-related attitudes gleaned through conversation) were also measured.

Many similar studies have found that the majority of online hostility will be committed by, accepted by, or even expected by those with higher levels of toxic online disinhibition—but as of yet, there has been relatively little research that inquires how civility can be encouraged in those predisposed to such toxic behaviors online. Notably, this study finds that those with higher levels of toxic, but not benign, online disinhibition prior to the task are capable of engaging amicably in deindividuated, anonymous online settings if their partner demonstrates a commitment to rationality even more than civility or intelligence, and at the same time does not compromise too easily.

CHAPTER 1.

INTRODUCTION

“Don’t read the comments!” has become an axiomatic law of life online, referring to the frustration Internet users might feel upon perusing the seemingly fruitless arguments other users have in the “comments” section of news websites, YouTube videos, Facebook posts, blogs, and other online spaces where audience feedback and conversation has been enabled. The inherent assumption is that arguments taking place in such spaces will not reach any meaningful conclusion, and that any expression of opinion will be met with hostility—engaging in such arguments is cautioned against: “Don’t feed the trolls.”

However, a large amount of political discourse has been displaced from face-to-face occasions into online environments—so the public’s aversion to deliberation in the online environment may be an impediment to political discourse occurring *at all* (Halpern & Gibbs, 2012; Levine, 2000). While it has been argued that the Internet might allow for greater democratic involvement by exposing people to opinions other than the ones they already hold, cynicism concerning the fruitfulness of such discourse may deter attempts at deliberative interaction. Incidental exposure, then, may be one of the most effective ways to encourage interaction between oneself and one’s political outgroup (Papacharissi, 2002; Wojcieszak & Mutz, 2009). Within more private forms of online communication (e.g., email, chat) there may be less risk of incidental exposure to debate than on social networking sites such as Facebook (Hermida, 2016), or on forums and microblogging sites such as Reddit, Tumblr, and Twitter (Cunningham, Nichols, Hinze, & Bowen, 2015). However, these spaces are increasingly becoming victim to the same

pessimistic outlook levied upon the dreaded comments section, with many people now expecting such environments to exacerbate offline conflicts or even give rise to new ones (Fox & Moreland, 2015). Essentially, then, every online space where deliberation could occur also becomes a place we are told not to waste our efforts, neutralizing the qualities of the online environment that would allow conversation to transcend the boundaries of time and geography, social anxiety, and other impediments to in-person debate.

Even outside of the academy, writers have pushed back against society's quick acceptance of this limiting philosophy of conflict avoidance, arguing that it enables abuse, stanches conversation, and trains us to believe that arguments—even arguments we would say are worth having—will not be worth the exhaustion we incur in the process (Dash, 2016; D'Kosta, 2013). Within the field of communication, especially as media psychologists, it is important to remember that while much research tends to focus on the anti-social effects of conflict that can be magnified by computer-mediated communication (CMC), not all CMC results in shallower, more toxic, or more dangerous communication. Rather, the digital media environment is merely a space in which “the separation of presence from communication” (Baym, 2010a, p. 4) significantly alters our relationships online, sometimes for the better.

At the individual level online, one can explore facets of their identity that would otherwise be constrained by temporal or geographical limits, as well as the pressures of face-to-face (FtF) communication. Individuals with high self-doubt, neuroticism, and emotional instability offline may feel safer in seeking social support online, and individuals who are more introverted can explore facets of their identities without the demands of in-person social interaction (Michikyan, Subrahmanyam, & Dennis, 2014). In

games and virtual environments, users can create avatars who are strong, successful, and look the way they would like to look in real life (Behm-Morawitz, 2013; Bessière, Seay, & Kiesler, 2007). Team-based projects incorporating individuals from many different countries at the same time can be accomplished, vastly extending the reach and efficacy of global economic, educational, and political communication.

Despite these lifted restrictions, scholarly optimism should be balanced with healthy skepticism; the unique challenges of online settings should not be ignored. Montoya-Weiss, Massey, and Song (2001) observed that international virtual teams must engage in more finely-coordinated collaborative conflict management in order to overcome miscommunication and other pitfalls of a reduced-cues environment. Valkenburg and Peter (2011) also note that the same elements of CMC that can promote self-esteem and foster stronger relationship quality among adolescent users of digital media can also promote cyberbullying and other anti-social behaviors. That is, there are not mutually exclusive “good” and “bad” aspects of digital media use—CMC affords greater control over self-presentation, and less perceived risk in self-disclosure due to features such as anonymity, asynchronicity, and remote accessibility.

Bailenson and colleagues (2008) describe a process of “transformed social interaction” (TSI) where such features of the online environment, which are often impossible to experience in FtF communication, produce new norms for social interaction, some of which may affect and be affected by the offline personality. For example, active Facebook use has been linked to higher extraversion and less social loneliness, but also less conscientiousness and higher narcissism (Ryan & Xenos, 2011). The social needs of an individual may determine the extent to which these features are

used for prosocial or anti-social behaviors. However, the amount of research featuring prosocial findings becomes more and more limited as the scope of analysis moves outward from effects of the online environment upon the individual, to an individual's real-life friends, to their extended social network, and finally to their interactions with anonymous strangers. Many scholars may reach the upsetting conclusion that, although anonymous and deindividuated online spaces are the exact ones that will naturally derive the most limitless, potentially constructive argumentation, they too are the ones that produce the most impediment to civil, deliberative progress. No user can be forced to shed their anonymity, nor should it be suggested that such an action would certainly alleviate online hostility. However, self-disclosure need not be limited to disclosure of personal identifiers—disclosure of one's perspective of the self, and of the other, may be an effective, humanizing strategy to engender civility in online deliberation.

In the following proposal, I will first describe the online environment and its barriers to civil communication, as well as give an updated review of research that has attempted, with varying success, to counteract those barriers. I will then argue for the undoing of recent cultural norms that have, in response to these barriers, labeled social media as a space in which fruitful debate is not achievable. Here, I will describe the migration of face-to-face communication to the online environment, and refer back to the dual-nature of computer-mediated communication attributes, which, far from producing barriers alone, can result in both positive and negative social outcomes depending on situational variables. Next, I will describe the additional problem of deindividuation and outgroup stereotyping, as well as the effects of emphasizing social, rather than personal, identities in online spaces. I will then define political deliberation as a necessarily civil

practice, setting up the main line of inquiry that this study seeks to address. Alluding back to the attributes of anonymous digital media that can generate incivility, I will then discuss argumentation patterns that may be employed to ensure that deliberation can occur in online task-based communication between interactants with polarizing social identities.

Although much online debate refers to specific policies and desired outcomes, I will explore whether individuals motivated by a goal of understanding their debate partner are more likely to experience reduced outgroup stereotyping and hostility, increased other-awareness (i.e., comprehension, if not appreciation, of another's viewpoint) and comprehension of opposing viewpoints, and increased chances of agreement upon the conclusion of a debate task—in other words, if such a goal can produce deliberation. An experimental procedure was employed to test the effectiveness of different deliberation frameworks—one that stresses a persuasive approach, and one that stresses understanding—on producing the aforementioned prosocial outcomes. Finally, implications for education will be discussed, namely the need for integration of civility practices in argument formation.

One goal of this dissertation will be to establish whether different argumentation frameworks can overcome states of disinhibition and ameliorate the conflict often inherent in anonymous encounters between opposing social groups—here, between Pro-Life and Pro-Choice individuals. By observing the qualities of deliberation that occur when working toward secondary goals related either to understanding or demonstration during a policy-related cooperation task, it may be possible to extract knowledge about online conflict that goes beyond attributes of communication platforms that lead to

disinhibition. Argumentation beliefs and styles will also be examined, including beliefs related to online disinhibition (both toxic and benign), attitude strength and level of awareness about one's own social groups. In response to the manipulated debate frames, it is expected that the successful completion of the task, perceptions of compromise, hostility, and other partner/opponent attributes, and resulting beliefs about the use of online debate in general, will shed light on whether stressing the importance of other-awareness during difficult conversations can make the experience more productive and less injurious to outgroup perceptions.

Beyond testing whether debate frames can be manipulated to encourage more prosocial interactions, this dissertation will reinforce existing literature that demonstrates the nature of digital space as one in which interactions between individuals will be loaded to carry intergroup implications more than offline communication might (Postmes, Spears, & Lea, 1998). That is, in reduced-cues spaces as most online spaces are, individuals, whether wittingly or not, will act as diplomats between the boundaries of social identities. You are never just “you” online. Depending on which identities you demonstrate, you are an example to others of, say, a Democrat or a Republican, a man or a woman, a religious or secular person. Of course this is also true in offline settings, but especially when we are boiled down to our words alone, it may be more likely that we demonstrate which “teams” we are a part of, rather than the things that make us unique—our personal characteristics that cannot be seen immediately through text, such as being a good friend or a hard worker.

Once we are categorized by others—and perhaps even by ourselves—we represent these categories through what we say and how we say it. Building upon

Postmes, Spears, and Lea (1998), this dissertation asserts that outgroup stereotypes are built through collective experiences of outgroup contact within deindividuated settings. In addition to cultural stereotypes built offline, stereotypes are constructed based on how groups tend to interact with one another online. If we are hostile, then, we may inadvertently contribute to the perception that the groups we represent are also hostile. In this study, Pro-Life and Pro-Choice interactants were examined to determine how contact with outgroup members or other ingroup members can affect stereotypes about these identities, and about how these identities communicate.

It may be considered desirable to take advantage of the disinhibited state, as I will discuss further, to advance difficult conversations—so discouraging toxic behaviors that may stem from this same process of disinhibition without damaging the disinhibited state entirely is the tightrope I will attempt to walk in this dissertation. Aspects of self-awareness may be key to this equation, if self-awareness can be considered indicative of more introspection upon one's actions online as well. As this is yet unknown, this study parses out elements of self-awareness as they relate not only to one's immediate actions and surroundings, but also to their presence in the online environment, and their membership to various social groups. It is unclear whether the majority of Internet users are aware of their contributions to perceptions of their various group identities, so in this study I will also be monitoring whether individual traits and states related to types of self-awareness are associated with benign or toxic disinhibition.

CHAPTER 2.

LITERATURE REVIEW

This literature review chapter will examine the history of group conflict and deliberation as they have evolved to adapt to the online environment, highlighting the need for more civil and effective online deliberation practices, and rejecting the notion that engagement in online argumentation is fruitless. The chapter will begin with an argument for the necessity of online deliberation, attempting to reconcile civility practices with the unique challenges of the online environment by proposing a comparison of “persuasion-based” deliberation goals with “understanding-based” deliberation goals. Next, I will review early digital communication models, followed by an in-depth explication of the online disinhibition effect and its impact on conversations between social outgroups. Finally, the review will conclude with an examination of the reduced-cues environment on personal identity and social identity, paying specific attention to the effects of deindividuation on the receiver of messages from outgroup members as they may be affected by various states and circumstances.

Political Deliberation as a Necessarily Civil Practice

Political deliberation itself has been defined differently by scholars, but one of the most frequently-cited definitions is that of Habermas (1989), wherein it was stated that deliberation must be: (1) Conversation between individuals who all have a stake in the solution of a political event or process, (2) who engage in the exchange of rational-critical arguments (3) with civility. Civility, he argues, is not a formality or nicety for the sake of avoiding hurt feelings. Rather, it is a necessary component of ethical deliberation, intrinsically connected to the exchange of rational-critical ideas. Turn-taking with and

active listening to one's opponent will allow a group of people to more effectively understand the needs of all group members, and allow deliberation to occur in such a way that will incorporate as much relevant information as possible—whereas hostility may suppress one's ability to incorporate the arguments of others into discussion of solutions.

Papacharissi (2004) later problematized earlier notions of hostility as a natural opposition to civility, arguing that scholars commonly conflate civility and politeness. Echoing Gurstein (1998), she writes that we must “reject a civility that leads to moral finger-wagging and an imposed sense of politeness, and strive for a sense of civility that is acceptable across cultural terrains” (p. 265). In this way, civility and politeness are delineated by a line separating that which is subjective and short-term (e.g., etiquette) from that which is a more all-encompassing drive for long-term protection of democratic society. Hostility in disinhibited settings has often been examined through evidence such as “flaming,” or the exchange of purposefully hurtful words (Gurak, 2001; O'Sullivan & Flanagan, 2003; Hutchens, Cicchirillo, & Hmielowski, 2014), and yet it is possible for people to say hurtful things to someone whose choices and autonomy are respected. Imagine a friend walking too close to the edge of a cliff, watching them almost fall, and shouting “You scared me half to death, you idiot!” Or, imagine an online conversation between two opposed individuals on the topic of trickle-down economics. While both individuals desire money to be distributed to lower-level employees, they disagree about how to accomplish this, and frustration may be exchanged—“That's never going to work, you dummy!”—while not disrespecting the value of the other's end goal. In this way, civil deliberation may occur even if it is impolite, and impoliteness may occur even if it does not result in incivility.

However, perceptions of impoliteness and incivility may be conflated during the practice of argumentation when conversation partners' arguments are rooted in different ontological or epistemological roots—especially in the reduced cues environment of CMC, or in a state of online disinhibition (Suler, 2004). Incivility may be perceived even in the absence of impoliteness. For example, consider an argument between two people, one secular and one non-secular, on the matter of abortion. With no impoliteness exchanged, each may perceive hostility in the other because the very fundamental nature of each one's position is a threat to the other. They may each, in different ways, perceive this hostility as incivility, because one perceives a threat to the soul, and one perceives a threat to human equality (Haidt, 2011). With each party arguing for an end-goal that will be irreconcilable with and preclude the values of the other (Gutmann & Thompson, 1990), each may believe that the other has malevolent, and therefore uncivil (Habermas, 1989), desires for their well-being. Although both parties may be acting ethically according to their own principles, the mere perception of hostility by one or both parties may be enough to hinder the deliberation. Some scholars have noted that perceptions of hostility may cause the termination of discussion, and that the prevalence of this phenomenon has led to wider spread computer-mediated communication apprehension and cynicism against online self-expression—that we must disengage with “trolls” at all costs, even if our hasty identification of “trolls” may cause us occasionally to shut down deliberation with a genuine and well-meaning opponent (Coles & West, 2016; Hunt, Atkin, & Krishnan, 2012).

Halpern and Gibbs (2013) later specified that, in addition to adhering to an ethical demeanor (e.g., using equality, symmetry, and civility in one's manner), logic and

rationality must be given priority over power and coercion for deliberation to occur. And yet, two opposing arguments may have both been constructed using logic and rationality, if they flow from differing ontological perspectives. The religious individual, who accepts a “truth” wherein the soul can be damaged by the acceptance of certain acts, may argue from this starting point in such a way that is linear and reasonable within this lens. The secular individual, who does not begin with belief in a divine moral code, similarly may argue in such a way that is linear and rational according to a foundation built on something other than belief in sin and soul. So, holding oneself to rigorous ethical and logical principles, which still stem from subjectivity no matter how logically applied, may not be a certain way to achieve deliberation.

Similarly, and more commonly, impoliteness may be perceived in the absence of incivility, and result in the same outcome as if incivility had occurred: The conversation becomes guarded, defensive, and potentially more polarization is perceived. Papacharissi (2004) found that impoliteness was more prevalent than incivility in a political subgroup of a popular forum, with 14.2% of messages containing incivility, and 22% containing impoliteness. As an example, name-calling may be considered a breach of etiquette, even if the message does not threaten to impede democracy or personal liberty—it can be impolite, and still not uncivil. There is a difference, then, between “you’re an idiot and you don’t deserve any rights,” and “you’re an idiot, but I respect your rights.” Name-calling may be especially rampant in scenarios where social identity is known but personal identity is unknown, increasing a reliance on out-group stereotypes. It is possible that this type of conflation of impoliteness with incivility occurs because one

party would presume the other's impolite treatment stems from an assumed intention to threaten their values.

Referring again to the conversation between the secular and non-secular individual, it is possible that the words "heathen" or "zealot" could be lobbed at one another, and without any explicit uncivil communication, this impoliteness would be taken, perhaps even correctly, as a cue that the other holds beliefs that would be threatening. Certainly, impoliteness and incivility may accompany one another often—but that does not mean that they refer to the same qualities of a conversation.

In addition to the subjective natures of ethics and logic, another threat to the success of deliberative communication may be found in power discrepancies between conversation partners. Hegemonic status can alter the extent to which deliberation is possible by introducing power and coercion into the arena, which Halpern and Gibbs (2013) argue is a direct impediment to the application of logic. Karpowitz, Mendelberg, and Shaker (2012) noted that previous literature may be too quick to assume that deliberation can be achieved similarly in any group arrangement, and that marginalized groups may be silenced, even when other group members are taking strides to avoid intentionally glossing over their needs. Using the example of women, the authors note that because women may be socialized to take a passive role in deliberation, men may wind up speaking more frequently or confidently, and therefore be disproportionately represented in a deliberative context. The authors suggest that in specifically solution-centric forms of deliberation, different rules should be applied to the outcome—in the event that men outnumber women, use a "unanimous" ruling, and in the event that women outnumber men, use a "majority" ruling. While this correction might seem as

though it would produce lopsided representation, it demonstrates that correction may be needed to preempt the inequalities produced by group differences in communication style (similar to an “affirmative action” type correction). Working at a more specifically political level, Ellis and Maoz (2007) found that, in online conversation between Palestinians and Israelis, Palestinians favored argumentation styles that promoted convergence and cooperation, whereas Israelis did not. Conceivably, this may be because Palestinians have more to gain by engaging in cooperation than Israelis, who currently occupy the (comparatively) hegemonic space and have no reason to engage Palestinians in deliberation about a change in the status quo. In this way, socio-cultural conditions may also determine the extent to which cognitive dissonance is resolved through motivated reasoning—or indeed avoidance of oppositional information altogether.

In one particularly optimistic piece, Kahan (2007) writes that although we may never be altruistic enough to escape the self-serving grasp of such biases, we may be able to mimic the conditions of altruism by finding a way to “erase the status threats that make people resist accurate information” (p. 154). This is an idealistic enough endeavor, to say nothing of the effects of confirmation bias that might result from unmotivated causes, such as cognitive limitation. What could possibly “erase” status threats? When the hegemonic status quo benefits only from one outcome, why would individuals in a position to protect such a status quo humor any discussion of alternative solutions that may benefit them less than another social group—especially in an online context, where it is so easy to avoid deliberation altogether?

Deliberation in the Online Environment

Political deliberation research obviously predates the existence of online spaces, beginning to gain popularity in the late 1980s, however, the field quickly readjusted to accommodate the spread of digital media in the late 90s and 2000s. Levine (2000) noted that more conversations, including political deliberation that had previously occurred in unmediated-world environments, were moving into the online space. This shift allows for discussion that had not before been able to occur due to geographic or temporal restrictions, or which would have previously been filtered through mass media renderings of other cultures and political landscapes. As previously mentioned, the lifting of such restrictions could be perceived as an attribute that would encourage global deliberation, or that the online environment would be a truly democratic “utopia.” However, Papacharissi (2002) later warned that scholars should avoid taking an extreme opinion about this shift—that it is an ultimately good or ultimately bad change—because online and physical world spaces are still parts of the overall public sphere, cyclically affecting one another. This cycle between online and offline behaviors is mirrored in the very style of communication inherent in the social media user’s experience, especially on such platforms as Facebook where users may know each other from in-person meetings, online discussions may be impacted by a user’s concerns about what future discussions could occur in offline spaces, although as described earlier, this would of course occur to a lesser extent on platforms where more and more of a user’s identity is obscured.

Current research maintains that social media use, specifically posting political opinions to social media platforms, is a strong causal factor related to subsequent voting behaviors (Settle et al., 2016). Despite this inherently social practice, where a user might

often craft such a post in full anticipation of the fact that they likely will receive feedback, deliberation is not guaranteed to occur, as many social networks are somewhat politically homogenous. For this reason, we may have a tendency to falsely remember a time before social media when selective exposure did not as strongly maroon us into such impermeable echo-chambers, when in actuality we have always had somewhat of an affinity to associate with like-minded people (Sears, 1967), and such polarization occurs generally in people who were already more heavily invested and polarized than most (Iyengar & Hahn, 2009).

Online Spaces as a Proponent of Confirmation Bias

Charles Mackay (1852/1932) famously stated that individuals with preconceived notions, rationally or irrationally held, will “torture facts into their service.” Referring to the practice of witch-hunting, Mackay observed that historical texts admit to even using different standards of evidence (often lower than that required for other crimes) to support accusations of witchcraft. While many historians now believe that the practice of witch-hunting may have had an intentionally-misleading, nefarious agenda—property acquisition—it still remains that for these accusations to effectively serve their purpose, there must have been a time when the majority of people thought that witchcraft was a believable accusation to levy upon a person. We in modern society might laugh about this, but the mechanisms that held such a belief in place may not be so different than mechanisms that hold other modern-day beliefs in place—for example, the power of prayer to heal a loved one, or the belief that one economic system is preferential to another. Lord, Ross, and Lepper (1979) noted that while many scholars previously believed that conservatives in particular, for whom “traditional” is a quality often

attributed, would be most susceptible to confirmation bias (see also: negativity effect), confirmation bias affects people from every ideology, and can be found on both sides of any belief. The authors note that in a test of individuals on both sides of the death penalty debate, supporters and denouncers alike would both quickly disparage evidence given in support of the opposing perspective.

Kunda (1990) surmised that such bias may actually be a form of motivated reasoning, wherein the attempt to reconcile cognitive dissonance caused by conflicting new information is somewhat of a conscious process. Nickerson (1998) later argued that confirmation bias should be defined by the possibility that there could be motivated or unmotivated methods of resolving cognitive dissonance. While motivated methods may seem much easier to understand at face value, we cannot rule out the possibility that cognitive limitations are responsible part of the time as well. Whichever the case, what we do find is that confirmation bias can be found even in modern day policy. Tuchman (1984) observes that policies which are adopted and integrated are much harder to unseat than for new policies to become adopted, because once a policy has gone through the trouble of becoming employed, subsequent discussions will rationalize the policy as though it is a constant—as though forgetting a time before the policy's implementation.

Empirical evidence of these phenomena can be found in just about any aspect of an individual's belief forming process. Within political evaluation, Lodge and Taber (2013) note that we tend to make automatic evaluations of a candidate or policy based on affective cues—for example, flashing the face of a happy person before asking questions about immigration caused participants to reflect more leniently upon illegal immigrants. The authors argue that voters make irrational choices, and then rationalize backwards to

justify their choices, even if the logic they cite was not present at the inception of their belief in a candidate's qualifications. Caplan (2007) previously noted that our social identities may also play a role in our initial political ideology formation—that we will do what our friends and families do, and as such we form unshakable “teams” built of like-minded people. Snyder and Cantor (1979) posit that the entire process of social evaluation may, to some extent, be built upon our satisfaction of preconceived notions of a person, rather than an impression of the person themselves. Kahan et al. (2012) found that even individuals gifted in numeracy are not exempt from confirmation bias associated with both sides of the climate change debate—in fact, they were possibly more affected by their social influences.

Still, Wilhelm (1998) did find that most political deliberation online only occurred in short bursts between like-minded individuals; arguably, this defeats one of the most useful functions of deliberation, which is to expose people to a plurality of perspectives so that a variety of solutions may be considered. Garrett, Carnahan, and Lynch (2013) later found that while selective exposure acts to strongly pull us toward like-minded online news sources, heavier use of attitude-consistent news sources does not correlate with also avoiding online news sources that feature opposing viewpoints (see also: Garrett, 2009). However, this effect may be specific to information sources that come in the form of news articles, rather than other people. Messing and Westwood (2014) found that within social media, distinctive selective exposure effects occur along the lines of social endorsement—that is, rather than relying on heuristics generated by attitudes about online news sources, instead we may be inclined to trust political information that has been approved by larger numbers of people; for example, a Facebook post with 12,000

“likes” may be more appealing to us than a post with only a hundred. As has been described using Social Judgment Theory (SJT), individuals are rarely capable of making decisions without some amount of social influence, as we tend to survey the range of possible attitudes and decide what the “extremes” are according to what attitudes are most popular. This influence may result in classic bandwagon effect, but on social media where our perception of society at large is strongly molded by the celebrities, friends, and politicians we follow, this bandwagon effect may resemble a bimodal distribution, where popular attitudes are split into polarized “camps.” Because the Internet allows us to access these social influences in higher volume more rapidly than before, it may then follow that, whether we intend to actively seek alternative viewpoints or not, our very positions embedded in social media networks of like-minded people will ensure that our incidental exposure to news media will be in line with our existing beliefs.

Arguably, the addition of outgroup-aversive tendencies become stronger when people freely discuss their explicit opinions, which may be more likely to happen outside of the constraints of journalistic objectivity (or whatever attempted semblance of objectivity may exist). To sidestep some of the more obvious clashes between political social identities, Wojcieszak and Mutz (2009) propose that incidental exposure to political topics may better foster deliberation than when politics are a central theme of the conversation. It is important to remember that online political information is not merely found in journal articles posted on social media. Rather, the conversations had between people commenting on the content (or, sometimes, perhaps just the title) also serve as information to the online bystander. Who posted the article? What do I know about their political beliefs? Is anyone combating these beliefs in the comments? What arguments

are they using? Which of these people am I closer to? The answers to these questions about the social nature of political information shared online are other sources of information that may be used when making decisions.

Demonstrating the Self, Deliberating with the Other

Kersting and Zimmerman (2014) note that merely increasing involvement in political discourse online may not lead to deliberative democracy, but rather to a network of isolated commentaries that do not invite conversation or debate, but are rather demonstrative in nature. The authors specify that the main goals of demonstrative participation in political dialogue are “to express political opinion and the belonging to a certain group” (p. 36), which signals closure to decisions having already been made, rather than ongoing political deliberation. This type of participation also reinforces ingroup/outgroup boundaries by demonstrating the range of stances acceptable in each group. Indeed this may be one of the largest differences between political talk and political deliberation; whereas talk requires no intention of reciprocity—similar to mass media—deliberation is an exchange (Calhoun, 1988; Mutz & Martin, 2001; Schudson, 1995). Specifically, deliberation in its ideal form is a civil, symmetrical exchange—talk has not been defined as such. Additionally, talk has no requirement of occurring exclusively between people whose personal fates may be tied to the outcome of a decision. On the other hand, both talk and deliberation may expose others, intentionally or incidentally, to a variety of perspectives, especially in an online environment where individuals may have access to a diversity of perspectives not immediately available to them in their unmediated life.

Certainly, both talk and deliberation may trigger attitude changes in others (Gastil, Black, & Moscovitz, 2008; Page, 1996), although persuasive language used in deliberative contexts must be rooted in logic and reason. So, the demonstration of one's social grouping would not qualify as deliberation, but perhaps the *justification* of one's social grouping might, if such a justification hinges on more than just affective or automatic, less rational processes. As noted, a great deal of political communication is demonstrative, not deliberative (Kersting & Zimmerman, 2014), and it may be the case that online communication does not *necessarily* change this impropotion, although we are exposed to more of each demonstrative and deliberative political communication. On the other hand, structural characteristics of many social media platforms, which encourage interactivity (Gane & Beer, 2008), may allow for more reciprocal exchanges to occur, and thus more productive deliberation. Whether this imbalance between demonstrative and deliberative practices has shifted because of digital media is yet unknown, however one purpose of this study will be to determine whether the disinhibitive (Suler, 2004) and deindividuating (Postmes, Spears, & Lea, 1998) characteristics of online spaces (to be described shortly) impede this potential. While this impediment to deliberation may exist, the increase in outgroup contact alone afforded by digital media spaces may affect deliberative processes.

One function of political discussion online with outgroup members may be to help individuals better understand which social labels they identify with most strongly (Gastil Black, & Moscovitz, 2008). This process also helps clarify to oneself what their stance might be in relation to a conversation partner's (Allport, 1954). This type of self-awareness may be a positive quality in conversation, if it results in desired outcomes such

as the accurate communication of one's needs. As established, deliberation requires reciprocal communication—where both parties are effective senders and receivers of information. It has been argued that partisan and deliberative practices during discussion will stem from different behaviors—either defending a team or considering all arguments fairly, which may then suggest that demonstrative discussion often precludes deliberation (MacKuen, Wolak, Keele, & Marcus, 2010).

It may take practice in “other-awareness” (i.e., a commitment to attempt understanding one's partner) to engage deliberatively. Son, Kornell, Finn, and Cantlon (2012) assert that self-awareness and other-awareness are types of metacognition that may have developed due to the competitive nature of being—mating, resource accumulation, and social protection all require us to figure out where we stand in the pecking order, and where others do, so that we can more easily identify allies and enemies. They write, “to thrive in an intensely social world requires humans (and perhaps other animals) to know themselves and to find ways to know the secret thoughts of others” (p. 171). Then, as metacognition would give rise to an increase in social wariness in an animal, we might find that empathy and other more altruistic traits would increase as well. We might especially expect that other-awareness in particular might allow for better understanding of a deliberation opponent's arguments, and less reliance on stereotypical information about their opponent's social group. As was found by Cameron and Trope (2004), we tend to assume that a political deliberation partner whose opinions are different than our own are in fact diametrically opposed to our own, and that the person represents a stereotypical opponent, deindividuated from the rest of their “camp.” They write that “prior expectancies should decrease as the amount of new individuating

information increases” (p. 669), meaning perhaps that more attentive interactants will be able to discern a deliberation partner’s specific views from the views of that person’s social group.

As Baym (2010a) noted, the reduced-cues environment provided by digital media creates conditions where, in lieu of personal identifiers, we may still make statements or associate with groups such that our social identities can become known to other people (see also: Lee & Pang, 2014). While this “demonstration” can aid us in finding allies, or maybe even precede more deliberative actions, it may also reinforce the echo chamber effect. As Postmes, Spears, and Lea (1998) have posited, online environments can reinforce boundaries between ingroups and outgroups. Because our social identities may become more immediately clear than our personal identities, when we engage with others online we may slip into a state of deindividuation. This deindividuation can affect many aspects of online communication. The diffusion of responsibility for hostility may remove inhibitions we have against reacting hurtfully to another, and we may prioritize “winning” over deliberation. Differences between ourselves and our conversational opponent may appear exaggerated or even hyperbolic, and we engage in more outgroup-stereotyping. For this reason, we might expect that individuals with particularly polarized (i.e., highly committed to one pole of the overall spectrum, individually) viewpoints will be more sensitive to identity threat, and thus perceive more hostility and engage in more stereotyping. Additionally, individuals in dyads with stronger differences between their attitudes may find it difficult not only to perform a deliberation task, but also to perceive civility and individuating characteristics within their opponents.

Referring again to the subject of confirmation bias, this time as may interact with threats perceived by outgroup member argument, Lepper (1979) found that strong negatively-valenced emotional responses—particularly anger—have the ability to magnify the biases one is likely to implement while searching for information or engaging in debate. It may then be that positive emotions—perhaps produced by different debate frameworks that accentuate commonalities between opponents—can also lessen tendency to rely on stereotypes or to resist cooperation.

Returning again to Halpern and Gibbs (2013), whose definition of deliberation was perhaps a clarified version of Habermas's original definition, we find that the tenets of deliberation do not perfectly map onto all of digital space in the same manner. The authors find that anonymous spaces are particularly prone to hostility. For example, Facebook may serve as a slightly healthier space for deliberation than YouTube, because people on Facebook may encounter more negative consequences for demonstrating hostility (their deliberative partners may be their friends and family, and their name and face will be associated with anything said, discouraging hostility). This is in alignment with literature related to cyberbullying, where aggressor-perceived anonymity has been linked with cyberbullying-supportive attitudes and behaviors (Barlett, Gentile, & Chew, 2016). So, the magic equation to find spaces for deliberation on the Internet may be high self-disclosure, and low deindividuation. However, as much of the Internet seems to lend itself well to exactly the opposite conditions—especially for political talk, which people are already quite squeamish about—scholars must continue to find other ways to encourage political deliberation within social media. Perhaps the first step in tempering the antisocial effects of anonymous, deindividuated social media without impeding the

positive qualities of such space may be as simple as incorporating more emphasis on attaining other-awareness in primary education. It is useful, then, to test whether individuals can be directed to engage in debate that would result in prosocial outcomes, more positive perceptions of their conversation partner, and cooperation. Thus, a first set of research questions is proposed to test whether instructing individuals to cooperate on a task and achieve an additional goal of either understanding their partner (an “other-aware” or “understanding” frame), or persuading their partner to understand them (a “demonstrative” frame) can achieve such outcomes:

RQ_{1a-g}: How will instructing participants to try to be “Other-Aware,” “Demonstrative,” or given no additional framework instructions during a deliberation task affect participants’ (a) engagement in outgroup-stereotyping, (b) identification with their partner, (c) cynicism following the task, (d) other-awareness, (e) perceptions of hostility, (f), post-test attitude change, and (g) perceptions of compromise?

With regard to the Halpern and Gibbs (2013) definition of political deliberation, it does not seem that deliberation can occur by simply going through the motions of ethical behavior—even if we are civil, engage in turn-taking and symmetry, there must be a concerted effort for each party to truly consider the perspectives of others, rather than become merely exposed to them. In addition to the emphasizing of commonalities that understanding-based deliberation frameworks may be able to achieve, it is important that deliberation represents a genuine desire to evaluate possible solutions, rather than just an exercise in demonstrating social grouping. In the section that follows, many of the concepts briefly touched upon—disinhibition, deindividuation, self-awareness, and

ingroup/outgroup construction—will be further described and set in concert with one another, prompting a number of additional questions about what qualities of online space and states of mind when engaging in such spaces are likelier to result in productive deliberation, despite the possible pitfalls of these spaces that must often be accepted alongside the positive characteristics.

Self-Disclosure (and Inviting That From Others) as an Antidote to Online Toxicity

Long before the advent of online communication, Gordon Allport (1954) theorized that mere contact with a social “other” may be enough to ameliorate group conflict, if that contact occurred under conditions of (1) equal status, (2) common goals, (3) cooperation between groups, and (4) authority support. This “contact-hypothesis” has been the basis for much optimism surrounding CMC’s ability to bring groups together for civil communication. Although not all models of CMC agree that attributes of the online environment coalesce to form a net-positive effect on relations between outgroups (see: social identity model of deindividuation effects, SIDE), some studies of online communication between conflicting groups have confirmed that Allport’s contact-hypothesis may still be relevant to digital communication. In a study of three Israeli populations—Jews, secular Jews, and Muslims—it was found that online contact was successful in reducing prejudice against cultural outgroups, especially in individuals whose initial opinions of their cultural other were particularly oppositional (Walther, Hoter, Ganayem, & Shonfeld, 2015). The Internet has the capacity to foster more intergroup contact in a smaller period of time than could ever occur in most FtF environments, and while it may immediately appear as though such interactions would necessarily be more brief or shallow in nature—perhaps less likely to result in

meaningful deliberation—there is evidence that social interactions have adapted to the online environment such that “deep” conversation can occur even between recent and anonymous acquaintances.

Walther (1996) originally posited that digital media allows for a more streamlined form of communication than FtF, creating an environment where interactions and relationships, even between anonymous parties, can become more intimate and direct much faster than before. This model, known as the hyperpersonal model, ushered in a new paradigm of digital media communication research that no longer assumed an impersonal nature of CMC—in fact, researchers now largely accept that CMC can exceed the advantages of traditional FtF features. Walther’s hyperpersonal model refers to four components of CMC: the receiver’s “idealized” perception of message senders, the message sender’s ability to optimize self-presentation cues, the reduced-cues and asynchronous nature of the online environment, and feedback processes unique to CMC which encourage behavioral confirmation—verification between sender and receiver that the influence is reciprocal.

Tidwell & Walther (2002) later specified that the veiled nature of CMC promotes the use of more direct uncertainty-reduction behaviors online, such as question-asking and increased/quicker self-disclosure. Self-disclosure, as described by Jourard (1971) long before the inception of social media, is “the act of making yourself manifest” to another (p. 19), and it has long been considered to be one marker of having a “healthy” personality (that is, a personality that allows for the successful integration of the self into one’s social setting).

Following the spread of the hyperpersonal model, digital media communication scholars began to focus more specifically on this element of rapid self-disclosure. In an analysis of five popular CMC theories, one common denominator that emerged was an agreement that CMC affords greater opportunities for increased breadth and depth of self-disclosure due to the compressed amount of time it takes to get through initial phases of trust-building that might take much longer in riskier FtF communication (Nguyen, Bin, & Campbell, 2012). Skeptics of the shift from FtF to largely digital communication may tend to view the CMC landscape from one of two extreme perspectives: (1) That we are our truest, purest, and most uninhibited selves online, or (2) That digital environments are populated by millions of false identities, lies, and exaggerations—nothing genuine. However, some of the earliest identity scholars have argued that there may not really even be one “true” self. Goffman (1959) described the “self” as a collection of identities, no single one any more “true” than another, no facet on its own able to claim representativeness of the entire person. Applied to digital media, this means that while self-disclosure in an online space may be incomplete, it can still be genuine (just as self-disclosure within a FtF setting is also incomplete but genuine).

Patterns of self-disclosure online seem to mimic offline communication styles. For example, Walton and Rice (2013) found that female Twitter users engaged in more frequent self-disclosure than male users, which mirrors traditional offline gender norms (e.g., communicative habits tend to be associated with femininity, whereas a reserved and impersonal demeanor is typically attributed to masculinity). This pattern may also be in alignment with notions of conflict-avoidance in women and assertiveness in men. Self-disclosure may be an important aspect of civil online conflict. At the intersection of self-

disclosure and gender, one might anecdotally remember the infamous #GamerGate ordeal (Parkin, 2014), wherein a large group of predominantly anonymous, male gamers defended and perpetrated the online harassment of women and feminist critics of the gaming industry. This conflict occurred almost completely in online spaces; perpetrators of harassment did not typically disclose personal information about themselves. These acts of hostility occurred specifically in environments that never required self-disclosure; more anonymous environments like Reddit, Tumblr, and 4chan played host to this altercation—and rarely spaces such as Facebook. In addition to anonymity, some online spaces (namely subreddits) where the bulk of hostility occurred were spaces in which offensive behavior was protected as “free speech,” reinforced by both community and moderators’ behaviors not only tolerating but rewarding the most outrageous posts with an algorithmic boost to the prized, highly-visible top of the page (Massanari, 2015). Traditional gender norms of communication persist into the online environment so it is expected that gender will play a major role in the study to follow, but it is important to examine other traits in individuals that impact communication styles, perceptions of others, and debate outcomes.

There is certainly evidence that antisocial behavior online may be strongly related to personality traits independent of online socialization messages—some going as far as to link antisocial “trolling” behaviors with sadism, Machiavellianism, and psychopathy (Buckels, Trapnell, & Paulhus, 2014)—or even that hostility and deception occur because it is assumed that “everyone does it” and it is enjoyable (Drouin, Miller, Wehle, & Hernandez, 2016). However, it has also been argued that transient states of mind, in concert with structural characteristics of social media platforms, can change the ways that

people behave online. In the following subsection, the online disinhibition effect will be described in greater detail, which pushes back against the notion that unchangeable psychological traits are responsible for behavioral changes in online settings.

Online Disinhibition

Scholars have long suggested that we feel the presence of our online conversation partners, even when using mediums that do not make use of real-time video chat. Speculating on a sense of togetherness that could occur in a virtual world divorced from the physical one, Goffman (1963) first coined the expression “co-presence” long before global penetration of online communication technology, referring to the phenomenon where people feel as though they are “accessible, available, and subject to one another,” which can occur even when geographically separated (p. 22). Although Goffman specifies that co-presence is more easily achieved when in physical proximity and when reciprocally felt, recent researchers have argued that mediated communication at a physical distance reduces, but does not fully impede, co-presence if indeed the functions of co-presence are to facilitate relationships, communities, and a sense of other-awareness (Nowak, 2001; Parks & Floyd, 1996; Parks & Roberts, 1997). Bregman and Haythornthwaite (2003) later noted that online co-presence is contingent upon the synchronicity of communication, much like traditional telephone, and that asynchronous forms of communication (e.g., blogs, microblogs, comment feeds) are much less likely to result in co-presence. The authors note that such asynchronicity and invisibility in anonymous online environments hastens online disinhibition. Online disinhibition refers to the phenomenon where online interactants feel more freely able to express themselves, less restrained by the pressures of FtF communication.

In his seminal paper discussing toxic online communication as it relates to the qualities of the online environment, Suler (2004) identified six factors of online communication that increase the potential for such a “disinhibited” effect: dissociative anonymity, asynchronicity, minimization of status and authority, dissociative imagination, solipsistic introjection, and invisibility. Suler cautioned against the assumption that these facets will always result in “worse” outcomes. He describes *benign disinhibition*, which he characterizes by a desire to share more openly, understand more fully, and engage with one’s conversation partner in a way that they may be inhibited from in a FtF setting. He also theorized that these facets could result in *toxic disinhibition*, which should be understood as resulting in more anti-social outcomes, such as aggression and incivility. For example, a person who feels more comfortable having a heated political argument on Facebook than in person may be experiencing online disinhibition—if the interactants use this state to feel more comfortable engaging in rude language and hostility, they may be enacting toxic disinhibition, but if they use their disinhibited state merely to share more openly how they feel, maintaining respect and civility, they may be enacting benign disinhibition. In the case of a debate between Pro-Life and Pro-Choice individuals, a benign effect may refer to each person feeling comfortable sharing their opposing views in a way they may be disinclined to do offline—this effect would then be considered toxic if civility breaks down and insults or other signs of disrespect displace friendly conversation. Phenomena such as cyberbullying (Barlett, Gentile, & Chew, 2014), sexting (Dir et al., 2013), and pornography consumption (Shim & Paul, 2014) may be influenced by a toxic disinhibition effect, but also complimenting and sharing praise on others’ photos and

posts (Suler, 2008), self-disclosure and other prosocial behaviors (Lapidot-Lefler & Barak, 2015) may be attributable to a benign disinhibition effect.

Although not all forms of digital media will promote all six factors of online disinhibition at the same time, some may commonly co-occur. For example, dissociative imagination particularly flourishes under anonymous conditions, and anonymity is much easier to accomplish under conditions of invisibility. However the facets relate, each of the six carries with it unique implications for online communication.

Dissociative anonymity. Described as “one of the principle factors that creates the online disinhibition effect” (Suler, 2004, p. 322), dissociative anonymity refers to the characteristics of online communication that make it possible for people to conceal or even change all or part of their offline identities. One’s online behavior can be almost entirely dissociated from their offline self, which allows for an avoidance of many types of negative consequences that would normally occur in response to problematic behavior. For example, an anonymous person who sexually harasses another online cannot as easily be reported to authority figures, and because their identity is unknown, they will not gain an unsavory reputation among friends, family, and employers. It has been found that news sites requiring commenters to provide their real names played host to less incivility than news sites allowing anonymous commenting (Fox, Cruz, & Lee, 2015; Santana, 2014). Still, anonymity should not be considered predictive of toxic disinhibition alone—for example, online survey data is often collected in an anonymous fashion so people will feel safer expressing their beliefs; similarly, anonymous online interactants may feel safer revealing their beliefs honestly and accurately, which is integral to discussion of one’s desires and needs.

Invisibility. Integral to how anonymity often comes into existence is invisibility, which refers to the inability for users to see one another. While there may be some visual presence of the self connected with our communication—profile pictures, for instance—with the exception of video chat, there is often no way to make eye contact with or read the body language of another. Anonymity and invisibility are similar in that they both involve the shrouding of the self—but whereas anonymity refers to the shrouding of identifying information, invisibility refers to the shrouding of the body and face. Other scholars have supported the isolation of these two constructs from one another, as one may be had without the other. Misoch (2015) described these same qualities as two subsets of the veiled identity, coining them instead as visual anonymity, where “the source of information cannot be detected physically,” and discursive anonymity, where “the messages of a person cannot be attributed to the sender” (p. 536). Suler (2004) writes that invisibility can have a disinhibiting effect because, just as we feel more comfortable having arguments or discussing personal information while avoiding eye contact during FtF communication, so too do we feel more comfortable discussing such topics online where we can be concealed entirely—and this benign comfort can, if not checked, give way to toxicity and hostility.

Specifically, invisibility may influence the extent to which individuals feel comfortable engaging in deception, such as tweaking one’s age or weight, but in some extreme cases more categorical identifiers such as ethnicity or gender (Warkentin, Woodworth, Hancock, & Cormier, 2010). In online dating settings, deception between invisible individuals—even individuals who plan to later meet in person—appears to occur along gendered lines; men tend to misrepresent their wealth, relationship

intentions, interests, and personal attributes, whereas women tend to misrepresent their weight (Hall, Park, Song, & Cody, 2010). This suggests that one's use of invisibility is dependent on context, and may change according to what characteristics an individual expects their conversation partner to prefer.

Asynchronicity. Asynchronous communication occurs when a dialogue is not executed in real-time. Email, YouTube comments, and forum posts are often composed in full awareness that one may not immediately receive a reply. This can be disinhibiting, Suler writes, because it allows people to avoid having to “cope with someone's immediate reaction” and instead put off the conversation until one is “willing and able to hear the response” (Suler, 2004, p. 323). Especially during online conflict when negative emotions are high, it may be tempting to engage in hostility and then “check out” of the conversation indefinitely, so as to avoid witnessing a reaction or prolonging the conversation. Other scholars have argued that instantaneous publishing may give rise to spontaneous types of hostility which, although sent quickly and to multiple receivers, can allow a person to asynchronously experience the ramifications of their words, if they do at all (Brown, 2017). While this dissertation is concerned primarily with synchronous online communication (e.g., chat), it is important to note that some effects of asynchronicity may still affect real-time textual communication. Compared with FtF or telephone communication, there is still a delay where one must compose a response—and this takes more effort than speaking aloud (Williams, 1977), so we may wish to say more in a single message than we would in a single turn speaking (Dennis, Fuller, & Valacich, 2008)—which may create anxiety in the other awaiting a response. Some messaging platforms such as Apple's iMessage have attempted to fill this void by implementing

what has been called a “typing awareness indicator” to let us know that our conversation partner’s response is on the way—however, some scholars have argued that this may increase anxiety, rather than assuage it (Puri & Mulay, 2016). This space between asynchronous and synchronous communication has the ability to create anxiety in interactants, especially during “high-stakes” discussions where we desire quick answers (Bennett, 2014).

Solipsistic introjection. Two important facets of solipsistic introjection can lead to a disinhibited effect, Suler argues. First, in the physical absence of one’s conversation partner, especially if that partner is a stranger, we often contrive a voice, face, and even a life that we imagine suits the type of person who would engage with us the way that they have. If the conversation is going well, that imagined other might be pleasant; if it is not, we may imagine them with an obnoxious personality, or an unappealing face and voice. Alternatively, we may hear their messages in our heads read with our own voice, which can make us feel as though the conversation is happening entirely within our intrapsychic space. Second, because this introjected persona is fleshed out in the context of our own imaginations, we may feel especially disinhibited because we feel as though we are having a conversation with ourselves. This combination of creating a “character” with assumed traits, and the enfolding of that character into our own very personal headspace can even create toxic disinhibition as it may cause us to project and impute unfairly on another during debate.

Very little research has isolated the process of solipsistic introjection, but it has been observed that solipsistic introjection can have an effect on online behaviors of ingroup members when conditions favorable for deindividuation are high. For example,

the Quran-burning incidents of 2010, which eventually led to violence in the nonmediated world, began with a construction of mob mentality spurred by solipsistic introjection experienced by Muslim and anti-Muslim Internet users. Al-Rawi (2016) observed that social media responses from supporters of Islam calling for action against the West tended to gain momentum whenever several similar opinions were voiced, perhaps because members of this ingroup began to mimic these behaviors, assuming that this type of online demonstration was desired by them as a member of the ingroup. At the heart of assumptions created by the solipsistic introjection process are beliefs held in one's own mind, and projected onto the online other. As is obvious in the case of outgroup members, these assumptions may be unfavorable—but even when interacting with ingroup members online, these assumptions may result in unfavorable outcomes, such as co-rumination. This is why, for the purposes of this study, hostility may not be predicted between outgroup members alone—two Pro-Choice interactants or two Pro-Life interactants may also feed off of one another and increase hostility, not against one another, but against the outgroup, in an attempt to demonstrate commonality.

Dissociative imagination. Just as we create a fantasy “other,” we may also create a fantasy self. As previously described, the shroud of anonymity may allow us to dissociate our online and offline selves. Beyond merely dissociating these two facets of our identities, we are also able to “be” someone else entirely online, if we should wish. In the context of games especially, we can imagine ourselves as great warriors or cunning spies, villains or royalty—any number of fantastic identities can be roleplayed in a way that allows our dissociated online self to evolve in ways that would be inaccessible to us in the real world. Even outside the context of explicit roleplaying and fantasy games,

online we can become the type of person whose traits we admire but are too shy in real life to emulate. However, this can result in toxic disinhibition when one feels as though they can displace responsibility for rude or abusive online actions onto this dissociated online self, thus disconnecting themselves not only from external consequences, but also personal feelings of guilt, regret, and responsibility. James (2014) later identified this effect as the “it’s just the Internet” mindset, where social media aggressors—especially within younger users—believe that their online selves inhabit a world where nothing “counts” and nothing said online should be taken seriously in the offline context (p. 93). In a focus group study of Facebook users, Fox and Moreland (2015) observed that “the very same participants who claimed Facebook was not ‘real’ all had stories to share about having negative emotions, offline conflict, or even relational termination in part because of (if not wholly attributable to) Facebook’s affordances” (p.173). So, while it may be tempting to believe that the online world allows us to inhabit an entirely separate self in an entirely separate place, believing in this false distinction may create conditions for toxic disinhibition.

Minimization of status and authority. Suler (2004) writes that online, “everyone—regardless of status, wealth, race, or gender—starts off on a level playing field, citing the “traditional Internet philosophy” of democracy and equal opportunity for expressing one’s voice (p. 324). While this may make immediate sense, given the removal of physical cues such as dress, posture, location, and the furnishings of one’s environment, it remains unclear to what extent offline status affects online status in ways that can be perceived by interactants. Suler concedes that “although one’s identity in the outside world ultimately may shape power in cyberspace, what mostly determines the

influence on others is one's skill in communicating (including writing skills), persistence, the quality of one's ideas, and technical know-how" (p. 324). In this way, Suler may be setting apart notions of offline power and persuasive skills, as though unrelated. These attributes may be interconnected, as evidenced by the persistence of a debater who is especially convinced of their own accuracy—often a confidence, even a default state of belief, reserved for those who occupy unchallenged hegemonic status. And yet, persistence and well-executed persuasive rhetoric will not always be indicative of a debater engaging in active listening.

Online Disinhibition, Self-Disclosure, and Attentive Listening

Online disinhibition does not result in uniformly “bad” communication—recall, the disinhibition effect can be benign (even “good”), or it can be toxic. Ho and McLeod (2008) found that people were more likely to express a minority opinion in a CMC condition than in a FtF condition, which they argue promotes public deliberation. On the other hand, Barlett, Gentile, and Chew (2016) found that aggressor-perceived anonymity was strongly associated with cyberbullying behavior. What this suggests, then, is that online disinhibition is not a phenomenon that occurs consistently across all users of a single media format. While there are certainly some elements of any given medium that are static—for example, a lack of real-time video chat on Twitter that renders users technically “invisible”—one cannot be said to be in a state of online disinhibition unless these technological traits are perceived by users in a way that would remove or temper inhibition. In this way, online disinhibition is a state contingent upon varying experiential qualities of media use—a subjective combination of user personality and motivation as they interact with fixed attributes of digital technologies.

One common denominator across Suler's six factors of online disinhibition is the removal of physical recognition one must have of a FtF conversational partner—in person, we cannot hide our faces, and we cannot impute a projected identity of the other. In fact, Lapidot-Lefler and Barak (2012) found that eye contact in particular was among the strongest aspects of visibility that could thwart negative effects of online disinhibition; thus, text-based forms of CMC could be particularly susceptible to toxic disinhibition, more so than video forms such as Skype and FaceTime. The authors later specified that while some online disinhibition effects may be avoided with the addition of eye contact, it is invisible, rather than visible, anonymous communication that results in the quickest self-disclosure of emotional states and details related to personal identity (Lapidot-Lefler & Barak, 2015). As described earlier, this means that the same aspects of online communication that may result in toxicity can also be responsible for the enhanced richness of social interaction.

Traditionally, in FtF communication, self-disclosure is logistically harder to avoid due to physical presence, although emotional self-disclosure may still be restrained and provide a barrier to meaningful debate. The consequences of uncivil discourse occurring in-person (e.g., damaged relationships, retaliation, tarnish of a personal reputation and so on) are more likely to be levied upon the uncivil individual whose name and face are known. CMC offers the ability for users to hide personal identifiers, shielding the perpetrator of online hostility from backlash that might occur in the physical world and enabling hostility at the same time that it enables self-disclosure, meaningful deliberation, and emotional exchange. Thus, it might be expected that benign and toxic processes of disinhibition will work independently. Benign disinhibition may encourage forms of self-

disclosure more easily met with attentive listening to a conversation partner, whereas toxic disinhibition may hinder the effects of self-disclosure, perhaps through hostility or assumptive listening. Thus, a first hypothesis is proposed:

H_{1a-b}: Those higher in (a) toxic online disinhibition, but not (b) benign online disinhibition will be worse at guessing how their conversation partner would have filled out an abortion attitude profile.

We do see that asynchronous media devoid of opportunities for eye-contact (Facebook, Twitter, forums like Reddit, even YouTube) are rife with contagious hostility—again, as evidenced by this unofficial, often-spoken, only half-joking rule of the Internet: “Don’t read the comments!” We especially observe this effect when cultural “hot topics” are brought up, such as politics, religion, and social justice (Gahagan, Vaterlaus, & Frost, 2016; Kwon & Gruzd, 2017). No longer do we look each other in the eye, but increasingly, we are also being encouraged to avoid looking at each other—especially at outgroup members—through written text online as well.

Online Disinhibition and Outgroup Contact

Despite our tendency to engage in selective exposure on and offline, outgroup contact occurs more often on the Internet than in the physical world (McKenna & Bargh, 2000). Scholars of the contact hypothesis—the notion that making acquaintances with outgroup members can reduce prejudice—have celebrated the potential for social media to create positive intergroup contact (Amichai-Hamburger & McKenna, 2006; Kim & Wojcieszak, 2018). However, this quality of the online environment may become dulled or even reversed under conditions of toxic disinhibition. If threatened, an ingroup member may engage in toxic communication habits not only because they believe it is

“safe” to do so (due to any of the six online qualities listed above) but also because their ability to listen attentively may become impaired (Inzlicht & Kang, 2010; Postmes, Spears, & Lea, 1998). In particular, anonymity and solipsistic introjection may lead an individual to make assumptions about another person’s character as a consequence of personal extrapolation from a limited amount of information gathered online. It may be difficult for individuals to avoid making such social judgments, as social cognitive processes may be triggered without conscious awareness, based on cues not consciously processed (Nisbett & Wilson, 1977).

Assumptions made about an outgroup member’s desires and conversational goals may produce an imagined argument and begin a cycle of mis/non-communication. For example, upon discovering that one’s online conversation partner is Pro-Choice, a Pro-Life individual may begin writing response crafted for what that person believes pro-Choice people usually resemble, making use of arguments that may not even be relevant to the specific conversation. The two may not even fully read one another’s replies before beginning to craft new responses, or reply only to the argument in parts (Rourke & Kanuka, 2007). In FtF communication, such interruption is easily witnessed and can then be addressed—the interrupted person may know to repeat what they had meant to say. On the other hand, CMC interactants may not know which parts of their argument went unheard.

A number of detriments may occur through toxic disinhibition—even persisting into offline spaces (Fox, Cruz, & Lee, 2015)—especially between outgroup members, aside from those mentioned above related to the synchronous aspects of the digital communication process (e.g., not reading an entire post before beginning to compose a

reply, arguing based on incomplete information). In order to preserve the qualities of the online environment that reduce barriers to controversial discussion, benign disinhibition should not be discouraged in the way we might discourage toxic disinhibition. When opposing social identities—even complex, political identities—are activated, it may be an automatic response to rely on stereotypes (Bargh, Chen, & Burrows, 1996; Iyengar & Westwood, 2015), which may result in frustration and toxic behaviors. Stopping this automatic response, then, might require some amount of self-awareness and restraint during one’s online interactions. Self-awareness is an important component in the “controlled” processes described in many dual-process theories, and may be able to combat toxic disinhibition and mistrust during outgroup contact (Hughes, Ambady, & Zaki, 2017; Lieberman, 2003). Here, self-awareness may manifest as more conscious introspection on one’s own behavioral tendencies in online spaces—self-awareness in a different context will be further described at a later point in this study. A second research question is proposed, to test whether disinhibition types are sensitive to self-awareness of online identity:

RQ₂: Will awareness of one’s own online presence be associated with either type of online disinhibition?

Additionally, it might be expected that, in part because of the aforementioned differences in socialized communication habits that exist between men and women, men would be more assertive in online conversations, perhaps relatedly engaging in more assumptive tendencies and quicker response formation. Not only did the #Gamergate phenomenon highlight the drawbacks of heated argument in anonymous spaces (Heron, Belford, & Goker, 2014); these drawbacks were strongly associated with social

identities—partially constructed through the opposing arguments themselves—those espousing critical feminism and those defending hegemonic masculinity in games communities (Chess & Shaw, 2015). These identities were strongly gendered, and thus the toxicity too contained highly gendered hostility, including stereotyping and rape threats, in addition to run-of-the-mill “make me a sammich” style jokes (Stein, 2016).

Although gender identity is certainly not perfectly mapped onto Pro-Life and Pro-Choice identities, nor is it the *only* identity related to the subject, it is nonetheless an identity strongly related to debate surrounding abortion. It may then make sense to expect similar toxicity—or beliefs supportive of hostility—along gendered lines in this study. This, in addition to the previously stated expectations we might find evident in men and women’s differently socialized communication habits, may then predict that men will have stronger beliefs consistent with toxic online disinhibition. Additionally—although this is so far based more in long-held perceptions of men and women than found by any widely agreed-upon metric—women are perceived by others to be better listeners (Broverman et al., 1972; Cohen, 2014) and more attuned to others’ emotions (Rosenkrantz et al., 1968) than men. If these perceptions of men’s and women’s listening capacity are at all related to information retention, it might be then expected that men will be worse at remembering what a conversation partner has said, or at being able to accurately piece together how their partner might have felt based off of clues given during a conversation. A second set of hypotheses is therefore proposed:

H_{2a-b}: Men will have (a) higher toxic online disinhibition scores than women will have, and (b) be worse at guessing how their conversation partner would have filled out an abortion questionnaire.

Beyond the disinhibition effect, as it has been briefly discussed, social identities salient to the topic at hand may be activated in defensive or offensive scenarios (Trepte & Loy, 2017; Turner, 1999). This replacement of the unique self with the social representative self may cause online interactants to quickly categorize each other as ally or foe. I now turn to a more explicit discussion of social identity theory, and how conflicts between members of differing social identities may magnify in online spaces.

Deindividuation and Hostility

Social identity theory posits that individuals make sense of their complicated social environments by constructing ingroups—groups that the individual feels they belong to—and outgroups—groups that the individual does not feel that they belong to (Tajfel, 1974; Tajfel & Turner, 1979). Ingroups may be constructed for substantive reasons such as sharing values, traits, and experiences with other members (or they may be arbitrary, a phenomenon known as minimal group paradigm), but ingroups are also defined to some extent by the othering of outgroups— people who share differing sets of values, traits, and experiences. Ingroups and outgroups are not all inherently diametrically opposed to one another, although belief that they are occasionally arises out of a process of ingroup favoritism. Ingroup favoritism, or the preference for individuals similar to oneself, can occur at a local level (e.g., cliques that form amongst school-age children) or at the global level (e.g., hegemonic structures that privilege whiteness and maleness), and often result in discrimination against or stereotyping of outgroup members (Brewer, 1999; Perdue, Dovidio, Gurtman, & Tyler, 1990).

Especially when *personal* identity is obscured, social media users may give off cues in text-based communication that emphasize their social identities, especially in the

context of debate between clear “camps,” like political or religious debate (Lea and Spears, 1992; Spears & Lea, 1992). On one hand, this can help us find like-minded individuals in communities where there can be a presumption of shared interests (Baym, 2010b). On the other hand, such cues can also alert others to our presence as part of an opposing group. Flanagin and colleagues (2002) found that women will actively work to maintain the reduced-cues nature of CMC so that their identity is not revealed and used to silence them in ways that occur offline. Especially when engaging in argumentation, and one’s very position may be enough to give away one’s social identity, social media users may desire a delicate balance of anonymity and personal differentiation from the stereotypes of their own social group—all while facing the same the temptation to lump their opponent a stereotyped social outgroup.

Hostility online can become even more exacerbated when individuals become “deindividuated.” Similar to anonymity and invisibility, this effect allows for the diffusion of responsibility in hostile online conflict. The social identity theory of deindividuation effects (SIDE) gives some insight into what occurs when opposing groups clash, especially in environments where individual personalities are suppressed in favor of the collective, causing outgroup opinions to be dismissed or oversimplified (Postmes, Spears, & Lea, 1998). This increase in cues related to social group, rather than individual identity, can strengthen the boundaries between social identities, causing sharper contrast and increasing the effects of outgroup-stereotyping (Kiesler, Siegel, & McGuire, 1984). Deindividuation, then, like the six characteristics of the online disinhibition effect, can have a benign or toxic effect. In one study that examined the intersection of online disinhibition with different cyberbullying roles, it was found that

just as aggressor-perceived anonymity increased cyberbullying behaviors, assuming the role of an ingroup defender—even when not anonymous—was also associated with the disinhibition necessary to use aggression *against* aggression—aggression in defense of another, rather than offense (Moore, Nakano, Enomoto, & Suda, 2012).

Still, it cannot be expected that all ingroup/outgroup interactions will necessarily result in an equal amount or similar nature of reliance on stereotypes when interacting with one another—members of marginalized outgroups have historically been stereotyped more than the White, straight, Christian, Western male (Rendon, 1984). It is unclear, however, whether this hierarchical understanding of identity would apply to Pro-Choice or Pro-Life labels. While the Pro-Choice identity might be associated with women and women’s health, it is also the identity associated with presently-legal abortion policy. On the other hand, Pro-Life individuals may be associated with Christianity and traditional gender roles, while it is also the identity *not* associated with presently-legal abortion policy. Because these identities are comprised of identities of mixed hegemonic status, it may not be immediately clear which identity will be more stereotyped than the other, or, beyond caucused identities associated with each, what traits will be used to stereotype each in online settings. Thus, a research question is proposed to examine these differences:

RQ₃: Will there be any differences in the ways that Pro-Life and Pro-Choice individuals stereotype one another?

Just as the hyperpersonal model suggests we “get to the point” in our online relationships much faster than in FtF communication, the SIDE model suggests that, when acting from within a deindividuated identity, we will take cognitive shortcuts to

more quickly produce and rely on social categorization and ingroup/outgroup classification in online conflict. For example, a one-on-one conversation about the 2016 U.S. presidential election may become very personal very quickly between two strangers online. With the addition of social identity (as a Democrat, as a Republican, etc), strangers may become hostile more quickly because they are not acting as individuals with any personal obligation of civility to their opponent, but rather more like a soldier in an army. That is, personal identity can become blurred together with the collective identity of other “Democrats” or “Republicans.” To wit, this intensification of social identity can promote stereotyping of outgroup members, as well as a diffusion of responsibility for any hostility that occurs in the process of arguing with such an outgroup member. To briefly digress, this is one of many possible reasons that political communication scholars have been able to identify for the increasing perception that online political deliberation has become polarized (see: DiMaggio et al., 2001; Farrell, 2011; Hwang, Kim, & Huh, 2014). Referring again to the #GamerGate conflict, we can see that a diffused, anonymous identity can result in especially toxic communication.

There are traits of CMC that, in contrast to FtF communication, foster more direct communication—uninhibited, direct communication. However, this is not to suggest that the conditions of online space will invariably result in toxic communication habits—just that CMC may allow individuals to avoid some of the social-psychological influences that inhibit FtF deliberation, such as fear of facing an outgroup member in-person (Stephan & Stephan, 1984). Therefore, CMC can foster deliberation of difficult topics, in part *because* of the very qualities that also make it a place of potential conflict (Ho & McLeod, 2008). Engaging in CMC may expose us to a plurality of perspectives, which is

an important aspect of democracy (McKenna & Bargh, 2000). However, it is important for scholars to continue researching qualities of civil online debate when it occurs *within* anonymous, deindividuated settings, rather than focusing on civility occurring outside of such conditions. To examine whether an environment in which a disinhibited state is highly likely—paired with deindividuation effects that would likely result from assumed arguments made by a political other—has any effect on perceptions of hostility and, if so, these perceptions are constructed via perceptions of rationality, intelligence, and/or civility, two more sets of hypotheses are proposed:

H_{3a-b}: Those who are high in (a) toxic online disinhibition but (b) not benign online disinhibition will perceive their partners to be more hostile.

H_{4a-c}: (a) Perceived rationality, (b) perceived intelligence, and (c) perceived civility will mediate the relationship between toxic online disinhibition and perceptions of partner's hostility.

As described, physically-evident traits may be scrubbed from one's online presence, but ingroups and outgroups still form, even in anonymous online discussions, through argumentation and debate—especially in political conversation, where one's ingroup may be easily triangulated by observing the nature of their concerns (Chen & Ng, 2016). Political deliberation, then, cannot be disentangled from the tenets of social identity theory, or from the consequences of ingroup favoritism. Referring back to the aforementioned example of political polarization, the perceived stark delineation between parties may be exacerbated by ingroup favoritism, implying that groups are “opposites”—increasingly defined by their status as “not the other.” After a reciprocal communication process concludes, when asked to recall information about another, or

another's argument, outgroup members may be less equipped to retell the positions argued by one another.

For example, suppose a Pro-Life individual—let's call her Sarah—discusses abortion policy online with a Pro-Choice individual—we will call her Jane. If Sarah posts a message where, in the first couple of sentences, she discloses that she identifies as “Pro-Life,” this demonstrates her social identity not only as one who espouses anti-abortion policy, but also may indicate a likelihood of other identities stereotypically associated with being Pro-Life, such as being conservative or religious. If Jane has a high tendency for toxic disinhibition, she may begin prematurely crafting a response relying only on this first information, without reading the rest of Sarah's post. Her post may have continued, “I don't believe in God, but I still believe a fetus's life must be protected at all costs.” Jane may counter Sarah with arguments crafted for the argument she expected Sarah to make, rather than the argument Sarah actually made, resulting in miscommunication and poor recall of what Sarah's actual argument was. If Jane replies by posting, “I don't believe anyone's religious beliefs should infringe on what women do with their bodies,” the conversation has begun on an almost nonsensical trajectory.

In order to salvage the discussion, Sarah will then have to post corrective information, which, as researchers have found, is often ineffective and sometimes has a worsening effect on the incorrectly held notion, if the corrective information is even noticed at all (Nyhan & Reifler, 2015). Corrective information may be effective absent any social influence (see: Bode & Vraga, 2015), but volatile online discussions are often influenced by indicators of social support for some messages over others, through comments (Hinnant, Subramanian, & Young, 2016) and often by “likes” and “upvotes”

(Winter, Brückner, & Krämer, 2015). In a closed space such as a private Skype chat room, however, interactants may be able to follow up misunderstandings with helpful corrective information. Still, topics that activate social identities may allow the deindividuation process to occur even in one-on-one settings, mimicking the effects of social influence and dampening corrective or nuanced information. It is expected that this study will uphold previous literature demonstrating that individuals hold more favorable perceptions of ingroup members than outgroup members. A number of damages may result from the process of deindividuation as it may occur between outgroup members, including specific effects on the following perceptions:

H_{5a-i}: Participants matched with a partner of opposing abortion stance (an outgroup member) will perceive their partners as (a) less rational, (b) less intelligent, (c) less civil, and they will (d) feel colder, (e) identify with less, and come away with more feelings of (f) perceived hostility, (g) political deliberation cynicism, and achieve (h) less other-awareness and (i) less task success than participants matched with a partner of similar abortion stance (an ingroup member) will perceive their partners.

Just as it was necessary to ask whether Pro-Life and Pro-Choice individuals would stereotype one another differently, so too must it be asked whether the negative outcomes asked above would occur differently between Pro-Life and Pro-Choice individuals. What, if any, differences exist between these groups that would cause outgroup contact to yield different perceptions and behaviors? Referring back to the question of whether identifiers of the Pro-Life or the Pro-Choice label hold hegemonic privilege over the other, it is still unclear whether one group's unique feelings of

disadvantage are more strongly held than those feelings in their counterpart group. So, one group may feel colder toward another, or one may listen less attentively to the other, focusing instead on defensive posturing. Identifying these group differences is crucial to gaining insight beyond online ingroup-outgroup relations alone—being able to situate these findings in their respective cultural context may allow this study to provide a more nuanced template for understanding other online relations between discussants of complex identities.

RQ₄: Will there be differences between Pro-Life and Pro-Choice individuals on the effects found in H₅?

As previously mentioned, Gastil, Black, and Moscovitz (2008) suggested that a prominent function of deliberation may be to better sort people into their ideological preferences. They found that in FtF settings, deliberation served primarily to increase the extent to which people knew whether their values were aligned with more liberal or more conservative values. While this may have a polarizing effect, it still serves as a way to expose people to a diversity of opinions. However, participants of the aforementioned study still only felt as though they had engaged in more deliberation when there was less variance in the attitudes of the group. In an effort to assess whether this phenomenon occurs outside of FtF settings, it is important to assess the effects that greater or lesser difference between opposing attitudes may have on “successful” deliberation. Although it might be expected that higher awareness of the boundaries of one’s ingroups and outgroups, as well as attitude polarization, might decrease an individual’s willingness to be cooperative and therefore result in less success during a deliberation task, it may also be true that these traits function only as an indicator of how well-sorted an individual is—

that is, how well they understand what their political and philosophical preferences are, and whether they know which social group holds similar preferences. Instead of hindering deliberation success, it may instead serve to foster interest, and thereby higher attention and commitment.

Govern and Marsch (2001) argued that self-awareness is a transient state rather than a fixed trait, and should be examined in the situational context of the moment. Building upon the authors' original scale, this study examines whether awareness of one's own social groupings will have an effect on cooperation. Here, this type of self-awareness refers to the extent to which a person is confident that they know which social groupings they belong to. It is expected that an individual might be more or less aware of their social categories depending on the day. For example, if a person recently spent quality time with a group of similar-minded individuals, or received support from someone of the same gender or ethnicity in a relevant scenario, this person might be more aware of their social categories than someone who was not recently primed to think about which social categories they belong to. Again referring back to the SIDE model (Postmes, Spears, & Lea, 1998), it may be the case that individuals' ability to cooperate with ingroup or outgroup members would be affected by how well-sorted they believe themselves to be, and how deindividuated this makes them more likely to perceive their own behavior. It is unclear whether this type of self-awareness would cause a person to be more careful about their own behaviors, or—in the case of highly polarized individuals—more committed to “winning” a debate than cooperating. Thus, a fifth set of research questions is proposed:

RQ_{5a-b}: Will a (a) self- awareness of one's social grouping and (b) attitude polarization have an effect on task success?

While awareness of one's own social grouping may be related to attitude polarization because stereotypes of different social groups may rely on the extreme cases within these groups, it may also be the case that people who are highly aware of which social groups they are a part of will be more well-versed in the nuances of each group's latitude of acceptance. For example, one might be very certain that they identify with the label "Pro Choice" because they tend to share opinions with other Pro Choice identifiers most often, and at the same time that person may not have a strong commitment to voting for entirely nonrestrictive policies, or may even prefer some restrictions such as a ban on third-trimester pregnancy termination (Fiorina, Abrams, & Pope, 2005). For this reason, an exploratory question about social identity awareness and attitude polarization is posed:

RQ₆: Will self-awareness of one's social grouping correlate with more highly polarized opinions?

This perceived polarization of social groups, which has been shown to escalate in response to uncivil online discourse, (Kwang, Kim, & Huh, 2014), may indeed be a reciprocal process, whereby incivility creates "otherness," and "otherness" creates incivility. Then, to know civility, one must truly exercise genuine desire to understand their political "other."

CHAPTER 3.

METHOD

This experiment employs three conditions based on the type of task(s) performed by the participants. Dyads were randomly sorted into one of three conditions, each corresponding to a different type of debate framework they were instructed to use during a task. Participants cooperated with one another to establish policies regarding abortion rights in a hypothetical territory, and, in the case of the two treatment conditions, attempt to do so while either trying to understand their partner, or getting their partner to understand them. In the control condition, participants did not receive any framework through which to engage in the task. The conditions will henceforth be referred to as “Other Aware,” “Demonstrative,” and simply “control.” Following the task, participants completed individual questionnaires assessing their partner and the task.

Sample Collection

To determine sample size, a power analysis was conducted using G*Power. Using Cohen’s (1988) recommendations for ANOVA designs, for three groups and any number of covariates it was found that 390 cases would allow for detection of small (0.20) standardized effects in two-tailed analyses with an alpha set at $p < .05$. Because analysis of dyads is also desirable, this sample size would still allow for detection of medium (0.50) and large (0.80) effects after halving of the full sample. Upwards of 500 university students were recruited for this experiment, expecting some hemorrhage between recruitment and study completion. Hemorrhage rates, however, were above 60% - still, the resulting 182-participant sample was able to achieve the necessary power for most analyses. Participants were recruited from communication and journalism courses, and

included desirable courses listed as requirements for general education, resulting in a diverse mix of ages and majors.

Procedure

Pre-Test and Dyad Sorting. Upon consenting to participate in the study, subjects were emailed a pre-test survey containing a short questionnaire about abortion attitudes and online disinhibition tendencies. This topic was chosen because previous research has shown that abortion attitudes are polarizing, and relevant discussion between Pro-Life and Pro-Choice individuals tends to be morally oppositional (Haidt, 2001). Carlton, Nelson, and Coleman (2000) found that individuals with more extreme abortion attitudes demonstrated predictably strong commitment to the issue (i.e., it is influential to their voting decisions), and yet both poles demonstrate the same level of commitment to the issue—that is, there is no evidence to suggest that motivations inherent in pro-Life attitudes are more inherently motivational than motivations inherent in pro-Choice attitudes, and vice-versa. Pew Research Center estimates suggest that Missourians are also divided over the issue, with 45% of Missouri residents in favor of legalization in all cases, 50% in favor of full illegality, and only 5% unsure (“Views about abortion by state,” 2014). Hess and Rueb (2005) found that political party and religious affiliation were strong predictors of abortion attitudes in college students. Despite the common notion that college towns, even in conservative states, tend to be more liberal and irreligious environments, Pew Data (2014) also shows that the majority of Christians in Missouri are college-educated, suggesting that Missouri’s large public universities do not have a substantial secularizing effect on their students.

Participants were asked to choose on a binary level whether they are “pro-life” or “pro-choice,” and then the strength to which they hold this opinion on a scale from one to eight. They then answered a set of questions to assess their specific attitudes about abortion. Participants also complete a variation upon the online disinhibition questionnaire (Udris, 2014).

Upon completing this survey, participants were asked to choose a range of days and times that they would be available to come into the lab to complete the rest of the study. Participants were sorted into schedule-compatible dyads composed of either one Pro-Life participant and one Pro-Choice participant (i.e., a “mixed-stance” dyad), both Pro-Life, or both Pro-Choice” (i.e., “matched-stance” dyads). Efforts were made to have an even distribution of difference between pro-life/pro-choice opinion strength (i.e., low-low, high-high, and high-low, low-high). The dyad was invited to come into the lab on one of their mutually-available days. Participants never knew who they worked with, as the dyads were invited to work at computer stations facing away from one another, with multiple dyads run at a time, and led to believe that their partner was actually in an entirely separate lab. Before the treatment task began, each participant filled out an adaptation of the Situational Self-Awareness Scale (SSAS) created by Govern and Marsch (2001). All communication between participants occurred through an entirely anonymous chat.

Treatment. This experiment utilized three possible conditions per dyad. In all conditions, participants were instructed to suppose that they are creating legislature for a new territory that, so far, has zero established policies concerning abortion. Using the text-based chatroom components of Skype (with no video or audio), each pair had 15

minutes to agree upon legislation options from a list. In the Other Awareness condition, participants were asked to complete the online deliberation task while attempting to understand where their conversation partner is coming from, and to truly comprehend what they believe. In the Demonstrative condition, participants were asked to complete the online deliberation task while attempting to get their partner to understand where they themselves are coming from, and to get their partner to truly comprehend what they believe. The final condition is a control, where dyads completed the online deliberation task with no additional instructions.

During the deliberation period, participants reviewed a list of possible options for legislative rules (Table 1). The list, which includes 16 situations for which abortion should be considered legal or illegal, includes pregnancy milestones (e.g., “a heartbeat occurs” or “the end of the second trimester”) and circumstances (e.g., “in the case of rape” or “if prenatal testing reveals illness”) that are commonly discussed in public and legislative spheres. A 17th item allows participants to write in a policy if they prefer one that is not listed (however, participants did not generally make use of this option, except in the case of one participant who wrote in an already-available option by error).

The dyad was instructed to discuss, using the text-only chat function of Skype, which options they agreed upon for their territory’s new legislation. They may choose to use a sheet of paper to personally keep track of their agreements and disagreements, and they may check as many items as they can agree upon (it was explained that some options necessarily rule out other options, e.g. there cannot be allowance for cases of rape if “restricted in all cases” has been checked). Each pair received 15 minutes to discuss their list(s). This time period allowed for dyads to take approximately one minute to discuss

each item and decide whether to select or reject it—a task which likely did not take a full minute for every single item.

Table 1. Possible options for a territory's new law. Check all that apply.

- 1 Restricted in all cases, and criminalized
- 2 Restricted in all cases, but not criminalized
- 3 Allowed in cases of incest
- 4 Allowed in cases of rape
- 5 Allowed in cases of mother being too young to safely deliver
- 6 Allowed in cases of mother being too ill/weak to safely deliver
- 7 Allowed in cases where the mother can demonstrate financial need for abortion
- 8 Allowed in cases where prenatal testing demonstrates illness in fetus
- 9 Allowed until a heartbeat occurs
- 10 Allowed up until the end of the first trimester
- 11 Allowed up until brain function occurs
- 12 Allowed up until the end of the second trimester
- 13 Allowed in any case where both biological parents agree
- 14 Allowed in any case where the mother does not want to be a mother
- 15 Allowed in all cases except partial-birth
- 16 Allowed entirely
- 17 OTHER: Explain

Post-Test. Upon completion of the fifteen-minute period of chat-based deliberation, participants navigated to a questionnaire, where they were individually asked to choose, from the list, the chosen legislature they agreed upon as a pair. They were then asked to rate different aspects of their current mood, rate aspects of the conversation and their conversation partner, compare their conversation partner to what they perceive as the “average” outgroup member, rate their agreement with items pertaining to optimism/cynicism about political talk, and state, in their own words, what they believe their partner’s belief was about the topic (which were then compared with their partner’s own responses to an abortion attitude profile from the pre-test sorting questionnaire).

Measures: Independent variables and covariates

Aside from condition, each individual participant's demographic information such as age, ethnicity, political affiliation, and gender were collected. Single-items were used to measure pro-Life/pro-Choice affiliation, as well as strength of commitment to this affiliation. At the dyadic level, it was noted whether individuals were demographically similar or if there were any divergences. All scale reliabilities were checked twice—once in the pilot test, and once during the full study—and are reported in Chapter 4.

Abortion attitudes. Participants' pre-test responses to an Abortion Attitude Profile questionnaire was used to assess individual attitudes about abortion. The scale was constructed for this study specially, attempting to have a balanced mix of statements that Pro-Life individuals might agree with for various reasons, statements that Pro-Choice individuals might agree with for various reasons, and some neutral items that both types of individuals might agree with (but might suspect the other of not agreeing with). All items were answered using the response options "yes," "it depends," and "no" (see Table 2 for all items). The reason for using categorical measures here, rather than scaled agreement measures, was to ensure that this profile could be easily used for later use in a post-task assessment of participants' other-awareness, wherein members of each dyad were asked to complete this questionnaire as they imagine their partner would completed it. Keeping the response options relatively simple allowed for people to have a more reasonable chance of accurately representing their partner's beliefs than they would have if response items had been scaled.

Table 2. Abortion attitude profile

-
- 1 Women usually regret having an abortion.
 - 2 Abortion should be a last resort.
 - 3 Abortion is murder.
 - 4 I would not mind if I found out a close friend had an abortion.
 - 5 A fetus is a living human.
 - 6 Abortion can provide relief.
 - 7 Abortion is a sin.
 - 8 Forcing women to carry an unwanted pregnancy to term is a form of slavery.
 - 9 There are no circumstances under which abortion is okay.
 - 10 Abortion is safe.
 - 11 There is dignity in motherhood.
 - 12 I believe women have the right to control their own bodies.

Online disinhibition. Participants' pre-test answers to the adaptation of Udris's (2014) online disinhibition scale were used to examine the extent to which participants' display trait disinhibition-prone beliefs about online communication. The final iteration of this scale can be found in Table 3. Items were measured in terms of the participants' agreement to each statement, using Likert scales ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree").

Table 3. Communication beliefs indicative of online disinhibition

Factor 1	<i>Benign disinhibition (Udris, 2014)</i>
1	It is easier to connect with people over computers or technology than in person.
2	The Internet is anonymous, so it feels easier for me to express my true feelings and thoughts online.
3	It is easier to write things online that would be hard to say in real life because you don't see the other's face.
4	It is easier to communicate online because you can reply any time you like.
5	I have an image of the other person in my head when I read their e-mail or messages online.
6	I feel like a different person online.
7	Online, I feel like I can communicate on the same level with others who are older or have higher status.
Factor 2	<i>Toxic disinhibition (Udris, 2014)</i>
8	I don't mind writing insulting things about others online, because it's anonymous.
9	It is easy to write insulting things online because there are no repercussions.
10	There are no rules online, so you can do whatever you want.
11	Writing insulting things online is not bullying.
Factor 3	<i>Solipsistic introjection</i>
12	I have a pretty good sense of who someone is, after I've read their posts online.
13	It is impossible to really know who someone is online, based off just a few posts.*
14	I have an easy time categorizing people after I read their posts online.
	* Denotes items which are reverse-coded

Opinion strength. Participants' pre-test responses were used to assess the strength with which they hold their opinion on the conversation topic on a scale from one to eight. At the dyadic level, the delta between the two participants' strength of opinion was measured.

Self-awareness. Participants' responses to an adaptation of the Situational Self-Awareness Scale (SSAS) created by Govern and Marsch (2001), taken immediately before the treatment task, were used to assess self-awareness. The original scale incorporates three factors—surroundings, public, and private—to assess an overall level

of self-awareness in the moment. An additional two factors were tested for the purposes of this study—social identity, and online identity. Participants were asked to rate the extent to which they agree with each item, using a seven-point Likert-type scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”). Refer to Table 4 for the full list of items and their factors.

Table 4. Adaptation of the Situational Self-Awareness Scale

Factor 1	<i>Self-awareness of surroundings (Govern and Marsch, 2001)</i>
1	Right now, I am keenly aware of everything in my environment.
2	Right now, I am conscious of what is going on around me.
3	Right now, I am conscious of all objects around me.
Factor 2	<i>Self-awareness of public self (Govern and Marsch, 2001)</i>
1	Right now, I am self-conscious about the way I look.
2	Right now, I am concerned about what other people think of me.
3	Right now, I am concerned about the way I present myself.
Factor 3	<i>Self-awareness of private self (Govern and Marsch, 2001)</i>
1	Right now, I am conscious of my inner feelings.
2	Right now, I am reflective about my life.
3	Right now, I am aware of my innermost thoughts.
Factor 4	<i>Self-awareness of social identity</i>
1	Right now, I am conscious of what social groups I belong to.
2	Right now, I am aware of what types of people are similar to me.
3	Right now, I am conscious of what I have in common with others.
4	Right now, I know who my social support is.
5	Right now, I am aware of which people I values interests with.
6	Right now, I know how to tell if someone else will be similar to me.
Factor 5	<i>Self-awareness of online identity</i>
1	Right now, I am aware of my online self.
2	Right now, I am conscious of how others perceive me online.
3	Right now, I am aware of how I am seen online.
4	Right now, I know how I behave online.
5	Right now, I am concerned with what other people will think of my online identity.
6	Right now, I am aware of how I tend to interact online.

Measures: Dependent

Task success. The instructed task was considered “successful” if both participants individually choose the same set of agreed-upon policies when asked following their deliberation. This variable, which was ultimately measured as presence or absence, and the strictest possible criteria were used to determine success: If members of the dyad were out of alignment by even one policy—for example, if one participant selected an option that the other left out—the task was considered unsuccessful.

Task compromise. Following the deliberation task, participants were asked to reflect upon the instances of compromise that occurred during the experience. Respondents were asked, “Who do you feel had to compromise most during the task?” and could choose between the options “Me,” “My partner,” or “Neither – we compromised equally.”

Mood and warmth. Participants were asked to complete an adaptation of the Positive and Negative Affect Schedule short form (PANAS-SF), developed by Kercher (1992). Instead of asking participants to rate the extent to which they have felt emotions over the past week, however, this adaptation asked them to rate the extent to which they are presently feeling the listed emotions. Responses are measured on a five-point scale, and the resulting data can be split by affect alone, or by parent-groups of emotions such as happiness, sadness, fear, anxiety, hopefulness and so on (Thompson, 2007; Watson, Clark, & Tellegen, 1988).

Additionally, participants responded to the question “How do you feel about your partner?” on a sliding scale from cold (0) to warm (100). Such scales as these, which have been popularized in political communication and affect research by Citrin and

colleagues' (1997) analysis of data from the American National Election Studies (ANES) have been reliably used to measure interplay between affect and social perceptions, as well as social distance.

Perceived Partner Hostility. Six items were used to rate the extent to which their partner seemed hostile. All items were created to test perceived hostility following a specific encounter with a specific partner, and the final two items were adapted from the verbal aggression and hostility factors of the Buss-Perry Aggression Questionnaire (BP-AQ) (Buss & Perry, 1992). See Table 5.

Table 5. Perceived partner hostility

- 1 My partner was never mean to me. (RC)
- 2 My partner was hostile.
- 3 I felt like my partner had a bad attitude.
- 4 My partner was nice. (RC)
- 5 I feel like my partner would make fun of people like me.
- 6 My partner seemed argumentative.

Perceived Partner Qualities: Nine items were used to measure the extent to which participants felt that the other members of the forum had been intelligent (intelligent, insightful, clever), rational (rational, reasonable, logical), and civil (civil, polite, considerate). Responses were collected using a seven-point Likert-type scale, from “Strongly disagree” to “Strongly agree.” Refer to Table 6 for the full list of items.

**Table 6. Perceived character qualities of conversation partner:
Intelligence, rationality, and civility**

- 1 My partner seemed civil
- 2 My partner seemed rational
- 3 My partner seemed intelligent
- 4 My partner seemed polite
- 5 My partner seemed logical
- 6 My partner seemed insightful
- 7 My partner seemed considerate
- 8 My partner seemed reasonable
- 9 My partner seemed clever

Identification with chat partner: Participants rated the extent to which they agree with statements about whether they related to and respected their conversation partner (e.g., “I’d like to be friends with someone like my conversation partner,” “My conversation partner was of good character”). This scale was loosely inspired by Hinkle and colleagues’ (1989) group identification scale, with items changed to refer more to one’s individual chat partner as a group representative, rather than the original scale’s references to fitting in with a collective. Also keeping in mind the effects of ingroup favoritism, this variation of the identification scale is meant to accomplish two simultaneous goals: (1) Examine the extent to which partners perceived themselves to be similar, and (2) examine the extent to which that similarity is rooted in a perception of sharing positive characteristics. A seven-point Likert scale was used, from “Strongly Disagree” to “Strongly Agree.” Exploratory factor analysis were used to pare down the scale. The full list of items can be found in Table 7.

Table 7. Items used to measure identification with chat partner

- 1 My conversation partner seemed like a good person.
- 2 I would like to spend time with people like my conversation partner.
- 3 I could not relate to my conversation partner.
- 4 I felt like I had a positive connection with my conversation partner.
- 5 I'd like to be friends with someone like my conversation partner.
- 6 I don't feel much respect for my conversation partner.
- 7 My conversation partner was of good character.

Online political deliberation cynicism. Participants were asked to describe how they feel about the worthiness of online political communication by responding to six statements on a scale from 1 (completely disagree) to 7 (completely agree). Because no cynicism scales presently exist to examine attitudes toward online political deliberation,

items were created to triangulate this construct. Exploratory factor analysis were used to pare down the scale. See Table 8 for full scale.

Table 8. Online political deliberation cynicism

- 1 Talking about politics online doesn't resolve much.
- 2 I enjoy talking about politics online with people of different opinions. (RC)
Sometimes it's not worth the exhaustion to debate with someone about politics
- 3 online.
- 4 The internet doesn't lend itself well to discussing politics.
- 5 Talking about politics on the internet doesn't do much except make people angry.
- 6 I think good discussions about politics can take place online. (RC)

Other-awareness. Participants were shown the Abortion Attitude Profile questionnaire from the pre-test phase and asked to attempt to “answer the questionnaire from their opponent’s perspective.” That is, they attempted to answer each of the 12 items (see Table 2) the way they believed that their opponent had answered. Each item that a participant was able to accurately match to their partner’s actual pre-test Abortion Attitude Questionnaire responses yielded an additional “point” toward this score. The extent to which their answers correspond to the answers given by their partner (their “points”) determined their “score” for this measure, with higher scores corresponding to more accuracy.

Outgroup-stereotyping. Participants were presented with a list of attributes, and asked to check whether they thought their partner could be described with any of them. Response options used to measure these items were “no,” “maybe,” and “yes.” See Table 9 for a full list of attributes. Higher scores on the index respective to one’s outgroup conversation partner was considered higher outgroup-stereotyping—participants’ perspectives on stereotypes related to their own ingroup were not considered because this measure is intended to examine only outgroup stereotyping. Because limited academic research exists on the stereotypes associated with Pro-Life and Pro-Choice identities

absent political identity, the attributes used here have been supplied using a thought-listing task provided in the pilot study (described more fully in a later section).

Table 9. Stereotypical traits associated with pro-life and pro-choice individuals

Factor	Attribute
Pro-Life	
1	Religious
2	Judgmental
3	Conservative
4	Unprogressive
5	Prudish
6	Sexist
7	Traditional
Pro-Choice	
1	Irresponsible
2	Promiscuous
3	Careless
4	Selfish
5	Liberal
6	Feminist
7	Sinful

Post-Test opinion strength. Finally, participants once more reported the strength with which they hold their opinion on the conversation topic on the same scale from one to eight.

Pilot Study

To test the efficacy of scales used in the full study, and to examine the landscape of political and abortion-specific attitudes in the population of interest, a pilot test was conducted using undergraduate students from an interdisciplinary communication course. A total of 144 students took part in the pilot study, and zero cases merited exclusion. The median age of participants was 19.91 ($min = 18$, $max = 26$, $SD = 1.33$), and the gender breakdown of the sample consisted of 41% men ($n = 59$), 58.3% women ($n = 84$), and

0.7% other ($n = 1$). Ethnicities represented in the sample included 85.4% White/Caucasian ($n = 123$), 6.3% Black/African-American/Afro-Caribbean ($n = 9$), 4.2% Asian/Asian-American ($n = 6$), 0.7% Middle Eastern/Middle Eastern American ($n = 1$), 0.7% Mexican-American/Hispanic/Latinx ($n = 1$), 0.7% Native American ($n = 1$), and 2.1% Other/Mixed Ethnicity ($n = 3$).

Participants' political leanings were relatively balanced, non-significantly leaning conservative ($M = 4.25$, $SD = 1.7$). The sample was comprised of 36.1% individuals ($n = 52$) who chose one of the three "liberal" options ("strongly liberal," "liberal," or "somewhat liberal), 14.58% individuals ($n = 21$) who chose "neither liberal nor conservative," and 49.31% individuals ($n = 71$) who chose one of the three conservative options ("strongly conservative," "conservative," or "somewhat conservative"). Religious affiliations represented in the sample included 40.3% Catholic ($n = 58$), 37.5% Non-Catholic Christian/Protestant ($n = 54$), 0.7% Jewish ($n = 1$), 1.4% Muslim ($n = 2$), 8.3% Agnostic ($n = 12$), 6.3% Atheist ($n = 9$), and 5.6% Other ($n = 8$). Strength of commitment to religious affiliation, examined on a three-point scale, was mid-range overall ($M = 2.23$, $SD = 0.72$). Because of the ordinal, perhaps subjectively interpreted nature of this scale, it was further examined categorically. Thus, a chi-square analysis revealed significant differences in commitment from religious groups $\chi^2(12, N = 144) = 26.54, p = .005$, *Cramer's V* = .304. The most striking of differences occurred between Catholics and Non-Catholic Christians, with Non-Christian Catholics reporting stronger identity with their religion than Catholics reported (see Table 11).

Table 11. Catholic v. Protestant Identity With Religion

	<i>How strongly do you affiliate with this religious affiliation?</i>			
	Not Very	Somewhat	Strongly	Total
Catholic	9	32	17	58
<i>Percent Catholics</i>	<i>15.52</i>	<i>55.17</i>	<i>29.31</i>	<i>100%</i>
Non-Catholic Christian	5	19	30	54
<i>Percent Non-Catholic Christians</i>	<i>9.26</i>	<i>35.19</i>	<i>55.56</i>	<i>100%</i>

Piloted Measures

A number of scales and measures were piloted using the sample described above. Abortion attitudes were first measured as a binary variable (options: “Pro-choice” and “Pro-life”), and then broken down into nuance. Participants who self-identified as Pro-Choice comprised 53.47% (n = 77) of the sample, and the remaining 46.53% (n = 67) self-identified as Pro-Life. It should be noted that there was some evidence that “pro-life” and “pro-choice” were not labels understood by the entirety of the sample – for example, there were some participants to labeled themselves “pro-life” and then gave attitudes about policy and the political outgroup that would be typical of someone who is “pro-choice,” and vice versa. Because of this, for the full study, a brief explanation of “pro-life” (one who prefers making abortion more illegal) and “pro-choice” (one who prefers keeping abortion more legal) preceded the question where participants were asked to place themselves on a binary. Responses to the eight-point attitude strength scale revealed slightly more variance in the attitude strength of Pro-Choice individuals. However, broken down into four categories, approximately equal groups were created: 27.1% (n = 39) reported being strongly Pro-Choice, 22.2% (n = 32) reported being weakly Pro-Choice, 25% (n = 36) reported being strongly Pro-Life, and 27.5% (n = 37)

reported being weakly Pro-Life. In some cases, it was necessary to split the full data file by binary response to this initial abortion stance item in order to better understand scales. When analysis using a split file occurs, it is noted in the sections below.

Outgroup stereotypes. As shown above in Table 9, stereotypes associated with Pro-Choice individuals and Pro-Life individuals were collected from the respective outgroup members of each group. This measure was adapted from thought-listing task procedures widely used in framing research (Price, Tewksbury, & Powers, 1997; Riles, Sangalang, Hurley, & Tewksbury, 2015), using a more thematic unpacking of the data rather than quantitative content analysis, to allow for guidance from a cursory examination of pop-journalism articles and books written by pro-Life and pro-Choice authors concerning stereotypes surrounding the perceived binary of abortion attitudes and arguments (Barklage, 2016; Callahan & Callahan, 1987; Douthat, 2013; Filipovic, 2014; Smith, 2016). In the pilot questionnaire, participants were presented with two open-ended, form-style pages on which to write the first five thoughts that come to mind when they think of Pro-Life and Pro-Choice individuals. Surmising from the participant's earlier response as to whether they generally identify with the Pro-Choice or Pro-Life label, the participant's ingroup and outgroup were noted, and responses to the thought-listing page for that individual's outgroup were collected. Thoughts relating to participants' ingroups were collected as well, but were not used for the purposes of this study.

The 144 participants in the piloted sample yielded 686 thought items (not all participants filled out all five thoughts for their outgroup members). Items were grouped together by how similar they were, most of which were duplicates—for example, many

people simply wrote “religious” as one of their thoughts listed for Pro-Life individuals. Each group was assessed for thematic similarity to other groups. While many people listed “religious” for Pro-Life individuals, some listed “Catholic” or “Evangelical” which could be considered thematically similar to “religious,” although neither of these more specific denominations occurred frequently enough to unseat the more general “religious” theme as being most prevalent. Seven themes comprised the vast majority of the thoughts relating to Pro-Life individuals—religious, judgmental, conservative, unprogressive, prudish, sexist, and traditional—with leftover items mainly referring to overly-specific personal anecdotes. Six themes comprised the vast majority of the thoughts relating to Pro-Choice individuals—Irresponsible, promiscuous, careless, selfish, feminist, and sinful. A seventh, “liberal,” did not occur with as much frequency as these themes, but was added as a counterpart to the “conservative” stereotype Pro-Choice individuals associated with Pro-Life individuals, and to balance out the number of stereotypes for each group. These items were then used in the full study using the aforementioned scale format, prompting participants to say whether they associated their conversation partner with any of the 14 stereotype items, using response items “no,” “maybe,” and “yes.”

Abortion attitude profile (AAP). The twelve items from the AAP were examined using a preliminary MANOVA and a subsequent series of univariate ANOVAs using a Bonferroni adjustment. Eleven items from the AAP yielded significant differences between Pro-Choice and Pro-Life responses, allowing for an understanding of typical responses from each group, and identifying one item that may be used to examine participants’ assumptions about their outgroup. The outstanding item, “There is dignity in motherhood,” was rated statistically similarly between Pro-Life and Pro-Choice groups.

Because of the ordinal nature of the response items, response frequencies for the options “false,” “it depends,” and “true” were also recorded. Refer to Table 12 for a full assessment of the AAP items.

Table 12. Responses to the Abortion Attitude Profile by Binary Abortion Stance

	Pro-Choice			Pro-Life			F-Test
	FALSE (1)	IT DEPENDS (2)	TRUE (3)	FALSE (1)	IT DEPENDS (2)	TRUE (3)	
1 Women usually regret having an abortion.	12 M = 1.90, SD = 0.45	61	4	1	40 M = 2.37, SD = 0.51	26 *	F(1, 142) = 35.25, p < .001, $\eta^2 = 0.20$
2 Abortion should be a last resort.	10 M = 2.30, SD = 0.69	34	33	7	3 M = 2.75, SD = 0.64	57 *	F(1, 142) = 16.22, p < .001, $\eta^2 = 0.10$
3 Abortion is murder.	46 M = 1.48, SD = 0.64	25	6	2	17 M = 2.69, SD = 0.53	48 *	F(1, 142) = 149.07, p < .001, $\eta^2 = 0.51$
4 I would not mind if I found out a close friend had an abortion.	8 M = 2.55, SD = 0.68	19	50	26	30 M = 1.78, SD = 0.71	11 *	F(1, 142) = 43.81, p < .001, $\eta^2 = 0.24$
5 A fetus is a living human.	28 M = 1.92, SD = 0.81	27	22	0	4 M = 2.94, SD = 0.24	63 *	F(1, 142) = 98.97, p < .001, $\eta^2 = 0.41$
6 Abortion can provide relief.	4 M = 2.62, SD = 0.59	21	52	21	26 M = 1.99, SD = 0.79	20 *	F(1, 142) = 30.91, p < .001, $\eta^2 = 0.18$
7 Abortion is a sin.	47 M = 1.51, SD = 0.72	19	10	3	14 M = 2.70, SD = 0.55	50 *	F(1, 141) = 120.14, p < .001, $\eta^2 = 0.46$
8 Forcing women to carry an unwanted pregnancy to term is a form of slavery.	20 M = 2.14, SD = 0.81	26	31	50	15 M = 1.28, SD = 0.52	2 *	F(1, 142) = 56.02, p < .001, $\eta^2 = 0.28$
9 There are no circumstances under which abortion is okay.	65 M = 1.16, SD = 0.37	12	0	19	35 M = 1.91, SD = 0.69	13 *	F(1, 142) = 69.65, p < .001, $\eta^2 = 0.33$
10 Abortion is safe.	4 M = 2.22, SD = 0.53	52	21	27	31 M = 1.73, SD = 0.69	9 *	F(1, 142) = 23.25, p < .001, $\eta^2 = 0.14$
11 There is dignity in motherhood.	3 M = 2.64, SD = 0.56	21	52	0	9 M = 2.87, SD = 0.34	58 ns	F(1, 141) = 7.86, p = .006, $\eta^2 = 0.053$
12 I believe women have the right to control their own bodies.	0 M = 2.94, SD = 0.25	5	72	8	34 M = 2.25, SD = 0.66	25 *	F(1, 142) = 70.79, p < .001, $\eta^2 = 0.33$

* = Results which are significant at the .001 level. Using a Bonferroni adjustment of .05/12, p-values must not exceed .004.

Abortion policy preferences. Abortion policy preferences were examined within pro-life and pro-choice binary groups. Again, it was explained to participants that choosing one of the extreme options (either “restricted in all cases and criminalized”/”restricted in all cases but not criminalized” or “allowed entirely”) would negate all options on the spectrum between. As might be expected by participants’ identity with their binary group, there were no “pro-choice” individuals who chose one of the “restricted in all cases” options, and there were no “pro-life” individuals who chose the “allowed entirely” option. A higher percentage of pro-life individuals chose their respective polarized preference than that of pro-choice individuals, who instead demonstrated greater variety of choices within the spectrum. Table 13 shows the breakdown of participants’ policy preferences by abortion attitude binary group.

Table 13. Possible options for a territory's new law. Check all that apply.

Item	Pro-Life (N = 67)		Pro-Choice (N = 77)	
	Poles (n = 40, 59.70%)	Excluding poles (n = 27, 40.30%)	Poles (n = 54, 70.13%)	Excluding poles (n = 23, 29.87%)
1 Restricted in all cases, and criminalized Restricted in all cases, but not 2 criminalized	11 (16.42%)	16 (23.88%)	0 (0%)	0 (0%)
3 Allowed in cases of incest		26 (38.81%)		42 (54.55%)
4 Allowed in cases of rape		36 (53.73%)		47 (61.04%)
5 Allowed in cases of mother being too young to safely deliver		22 (32.84%)		43 (55.84%)
6 ill/weak to safely deliver		31 (46.27%)		47 (61.04%)
7 demonstrate financial need for abortion		5 (7.46%)		33 (42.86%)
8 demonstrates illness in fetus		9 (13.43%)		35 (45.45%)
9 Allowed until a heartbeat occurs		7 (10.45%)		19 (24.68%)
10 Allowed up until the end of the first trimester		5 (7.46%)		20 (25.97%)
11 Allowed up until brain function occurs		2 (2.98%)		14 (18.18%)
12 Allowed up until the end of the second trimester		0 (0%)		8 (10.39%)
13 Allowed in any case where both biological parents agree		6 (9.00%)		23 (29.87%)
14 does not want to be a mother		0 (0%)		23 (29.87%)
15 Allowed in all cases except partial-birth		1 (1.49%)		11 (14.29%)
16 Allowed entirely	0 (0%)		23 (29.87%)	
17 OTHER: Explain		4		3

*Note = Percentages displayed in the "excluding poles" columns are taken from total group Ns

With regard to the cases in which participants selected “other,” they were given the option to write in what their specific opinion would be. Of pro-life participants who selected “other,” two wrote in that they would allow a termination procedure if the baby would be very likely to be delivered with major heart problems or brain-death. The other two wrote in that they were too ambivalent about cases of rape or cases where the mother’s life is threatened to commit to a stance. Of the “other” responses written in by pro-choice participants, two specified that if the mother “or father” does not want to be a parent then abortion should be allowed. It is not clear if these participants are asserting that a biological father should be able to mandate an abortion, or simply be enough to allow one. The last “other” written in by a pro-choice participant simply stated “I don’t care.” Again, it is not clear whether this apathy would qualify for “allowed in all cases.”

SSAS - surroundings, public self, and private self. To ensure that the original SSAS items proposed by Govern and Marsch (2001) would suit this particular population, reliability analysis was performed on each factor to determine whether each could be examined independently from the full scale, and once again for the full original three-factor scale. It is important to note that 65.3% (n = 94) of participants reported taking the questionnaire in a “mostly private” space, and the remaining 34.7% (n = 50) in a “mostly public” space, therefore the quality of privacy felt by pilot participants would be important to conserve for the full study. The “surroundings” factor yielded good reliability (Cronbach’s $\alpha = 0.833$), the “public self” factor also yielded good reliability (Cronbach’s $\alpha = 0.811$), and the “private self” factor yielded acceptable reliability (Cronbach’s $\alpha = 0.797$). The full scale yielded acceptable reliability (Cronbach’s $\alpha =$

0.712). The addition of the two proposed factors, outlined below, increased the full scale's reliability (Cronbach's $\alpha = 0.814$).

SSAS - social identity. Six items comprised the Social Identity factor of this study's extension of the Situational Self-Awareness Scale. Participants were asked to rate the extent to which they agreed with a statement pertaining to their confidence in and knowledge of their own social identity, on a scale from one to seven. Principal components factor analysis concluded that each of the six items (refer to Table 4) loaded onto a single factor, and so no items were excluded or separated. The scale yielded good acceptability (Cronbach's $\alpha = 0.809$). Participants in the sample were mostly confident in their social identity self-awareness ($M = 5.71$, $SD = 0.75$).

SSAS - online identity. Similarly, factor analysis using principal components analysis was used to determine whether each of the six proposed items (refer to Table 4) loaded onto a single factor. One item ("I am concerned with what other people will think of my online identity") did not fit with the five other items (factor loading = 0.111), and after theoretical examination of the variable, it was determined that this item conflated awareness of online identity with awareness of others' perceptions of one's online identity, so it was removed from the scale. The resulting scale yielded good reliability (Cronbach's $\alpha = 0.857$).

Online disinhibition – benign and toxic. Online disinhibition was measured using an expansion of the Udris (2014) Online Disinhibition Scale, which measures benign and toxic disinhibition. The original study, comprised of a high school sample from Japan, demonstrated subscale reliabilities of reaching $\alpha > 0.81$. Here, reliability for the Benign Disinhibition subscale was acceptable (Cronbach's $\alpha = 0.714$), and reliability

for the Toxic Disinhibition subscale was good (Cronbach's $\alpha = 0.827$). Total scale reliability was acceptable (Cronbach's $\alpha = 0.758$).

Online disinhibition – solipsistic introjection. The three items used to measure solipsistic introjection did not yield an acceptable reliability (Cronbach's $\alpha = 0.504$). Factor analysis revealed that responses to the reverse-coded second item (“It is impossible to really know who someone is online, based off just a few posts”) were widely variable, so it was thrown out. However, the resulting two-item scale still did not quite achieve acceptable reliability (Cronbach's $\alpha = 0.672$) and therefore this subscale was not retained for the full study.

Cynicism. Six items were included to assess online political deliberation cynicism. With all initial items included, the scale yielded acceptable reliability (Cronbach's $\alpha = 0.754$). However, factor analysis revealed that the second item, which was reverse-coded (“I enjoy talking about politics online with people of different opinions”), was the only item that did not load onto the principal factor. Theoretical reassessment of this item deemed it unlike the rest, in that the other five items had more to do with the usefulness of political deliberation online, whereas this item measured enjoyment only. The item was thrown out, and the resulting scale yielded better, still acceptable, reliability (Cronbach's $\alpha = 0.773$).

Additional observations. Pilot data supports claims in previous literature that suggest abortion attitudes are strongly linked with religion and political identity. Here, chi-square analyses reveal that abortion attitudes are related to religious preferences $\chi^2(6, N = 144) = 18.49, p = .008, Cramer's V = .358$, and to binary political preference (i.e.,

excluding “moderates”), $\chi^2(1, N = 123) = 34.72, p < .001, Cramer's V = .531$. Exact counts can be found in Table 14.

Table 14. Abortion attitude binary by religion and political identity

		Pro-Choice	Pro-Life	Total
What religious affiliation to you identify with?	Catholicism	23	35	58
	Non-Catholic			
	Christian	27	27	54
	Judaism	1	0	1
	Islam	1	1	2
	Agnosticism	11	1	12
	Atheism	7	2	9
	Other	7	1	8
Total		77	67	144
What is your political identity?	Liberal	44	8	52
	Conservative	22	49	71
	Total	66	57	123

Description of main study sample:

A total of 182 students, forming 91 dyads, took part in the full study. As in the pilot, zero cases merited exclusion. The median age of participants was 19.80 ($min = 18, max = 26, SD = 1.47$), and the gender breakdown of the sample consisted of 34.6% men ($n = 63$), 65.4% women ($n = 119$). Ethnicities represented in the sample included 80.2% White/Caucasian ($n = 146$), 8.8% Black/African-American/Afro-Caribbean ($n = 16$), 7.7% Asian/Asian-American ($n = 14$), 2.2% Mexican-American/Hispanic/Latinx ($n = 4$), and 1.1% Other/Mixed Ethnicity ($n = 2$). Middle Eastern/Middle Eastern American and Native American populations were not represented in the full study.

Participants' political leanings were relatively balanced, non-significantly leaning liberal ($M = 3.75, SD = 1.62$). The sample was comprised of 43.4% individuals ($n = 79$) who chose one of the three “liberal” options (“strongly liberal,” “liberal,” or “somewhat

liberal), 14.8% individuals ($n = 27$) who chose “neither liberal nor conservative,” and 31.2% individuals ($n = 57$) who chose one of the three conservative options (“strongly conservative,” “conservative,” or “somewhat conservative”). Religious affiliations represented in the sample included 28% Catholic ($n = 51$), 40.1% Non-Catholic Christian/Protestant ($n = 73$), 4.9% Jewish ($n = 9$), 0.5% Buddhist ($n = 1$), 0.5% Hindu ($n = 1$), 8.8% Agnostic ($n = 16$), 6.3% Atheist ($n = 9$), and 12.1% Other ($n = 22$). There were no Muslim participants in the full study, and it should be noted that Protestants displaced Catholics as the predominant religious affiliation, in comparison to representations in the pilot. Strength of commitment to religious affiliation, examined on a three-point scale, was mid-range overall ($M = 2.12$, $SD = 0.76$). Again, a chi-square analysis revealed significant differences in commitment from religious groups $\chi^2(14, N = 182) = 24.77$, $p = .017$, *Cramer’s V* = .312. Again, Non-Christian Catholics reported stronger identity with their religion than Catholics reported. Considering Protestants’ higher identity with their religion than that of Catholics, their more dominant representation in the full study than in the pilot may be expected to alter some of the findings.

Finally, the full study sample leaned more strongly toward Pro-Choice attitudes than the pilot study, with 71.4% ($n = 130$) of participants identifying as “Pro-Choice,” and the remaining 28.6% ($n = 52$) identifying as “Pro-Life.” An attempt was made to keep the proportion of Pro-Life and Pro-Choice individuals in matched-stance or mixed-stance dyads similar (via randomization), but the results of this attempt were slightly skewed such that a significantly larger proportion of Pro-Life participants (69.23%, $n = 36$) were placed in mixed-stance dyads, whereas only 27.7% ($n = 36$) were placed in

mixed-stance dyads. Wherever possible, this imbalance is examined—quite often in the analyses that follow, either stance identity or dyad composition is either controlled for, or examined specifically as a variable of interest.

CHAPTER 4.

RESULTS

The following chapter presents the results of the hypotheses and research questions. Primarily, group differences were assessed using ANOVA, MANOVA, and chi-square procedures, pre-test measures predictive of post-test outcomes used regression analyses, and associations between measures embedded in the same chronological segments of the procedure were tested using bivariate correlation analyses. As will become evident, some exploratory follow-up analyses occur for the sake of gathering additional information; however, the inductive nature of these exploratory analyses means that future studies may consider attempting replication before retroactively fitting theoretical implications to what is found here.

Although this will be more thoroughly described in the analysis of Research Question 1, it is important to note up-front that the debate frame conditions did not have any effects on any of the other variables, nor were they, even accidentally, associated with any participant traits measured in the pre-test. Thus, the conditions were initially used as a control variable in all other analyses to be absolute certain that they did not magnify or dampen any other effects used in this study—but it was found that controlling for condition never altered the significance of any significant effects or null findings, so the following write-up of these analyses uses models where this control has been pruned to conserve power. Additionally, efforts to ensure there was no lopsidedness in the distribution of matched/unmatched stance dyads in the conditions (randomization) indeed succeeded in producing categories with no over or under-representation of any dyadic composition (both Pro-Life, both Pro-Choice, one Pro-Life and one Pro-Choice) within

the debate frame conditions (Other Aware, Demonstrative, Control), $\chi^2(4, N = 182) = 5.29, p = .259$. There is, therefore, no relationship between dyad composition and condition, and therefore they will be treated as independent variables.

Research Question 1a-g:

The first set of research questions examined how different deliberative frameworks might affect individuals' perceptions of the experience and conversation partner, including (a) engagement in outgroup stereotyping, (b) identification with partner, (c) political deliberation cynicism, (d) other-awareness, (e) perceptions of hostility, (f) attitude change, and (g) perceptions of compromise. In this study, participants were fairly evenly divided into the three conditions, with 36.3% ($n = 66$, 33 dyads) in the control condition, 29.7% ($n = 54$, 27 dyads) in the Other Aware condition, and 34% ($n = 62$, 31 dyads) in the Demonstrative condition.

To examine RQ_{1a-e}, a MANOVA was performed to simultaneously test for a batch of similarly measured post-test scales. Results of the MANOVA indicated that there was not a significant effect of condition on stereotyping, identification with partner, political deliberation cynicism, other-awareness, or perceptions of hostility, $F(2, 346) = .990$, $Wilks' \lambda = .955, p = .444$. In turn, univariate tests for these five variables were nonsignificant.

To test RQ_{1f}, a repeated-measures ANOVA was used to determine whether condition affected the difference between pre-test and post-test abortion attitude scores. No differences were detected in the Δ between pre and post test scores within the control, Demonstrative, and Other Aware conditions $F(2, 179) = 1.39, p = .251$, observed power = .297.

However, when examining the impact of condition on attitude change within dyads comprised of either two pro-Life, two pro-Choice, or one pro-Choice and one pro-Life individual, it was found that within dyads comprised of two pro-choice individuals, there was an effect in the control group such that post-test stance toward abortion became more strongly pro-Choice after the treatment $F(2, 93) = 11.99, p < .05 (p = .044), \eta^2 = .065$. That is, there may be evidence of a natural (that is, when given no instructions about how they should engage) tendency for pro-Choice individuals to further polarize when discussing online with someone of the same abortion stance. The same effect did not occur in individuals in dyads with two pro-Life participants, or in dyads with one pro-Choice participant and one pro-Life participant.

Finally, to test whether the conditions had any effect on perceptions of compromise (RQ_{1g}), a chi-square analysis was conducted. Again, respondents were asked, “Who do you feel had to compromise most during the task?” and could choose between the options “Me,” “My partner,” or “Neither – we compromised equally.” Results of the analysis showed no significant differences in perceptions of compromise within the conditions, $\chi^2(4, N = 182) = 3.97, p = .410$. For the sake of gaining additional insight into the precursors of perceived compromise, a MANOVA was conducted (See Table 15) to assess differences in other deliberation qualities according to perceptions of compromise.

Table 15. Perceptions of one's conversation partner and perceptions of compromise.

	Me	My Partner	Neither - we compromised equally
1 Rationality	5.60 (1.32) ^{a***}	5.75 (1.09) ^{b*}	6.34 (0.73) ^{a***b*}
2 Civility	5.83 (1.02) ^{c**}	6.02 (0.90)	6.35 (0.74) ^{c**}
3 Intelligence Feeling	5.11 (1.52) ^{d**}	5.29 (1.18)	5.85 (1.05) ^{d**}
4 Thermometer Identification	67.36 (22.1) ^{e***}	68.9 (19.46) ^{f***}	84.74 (14.7) ^{e***f***}
5 with Perceptions of	4.96 (0.99) ^{g***}	5.34 (0.85) ^{h*}	5.82 (0.86) ^{g***h*}
6 hostility Other-	2.11 (1.15) ^{i*}	1.87 (0.77)	1.65 (0.86) ^{i*}
7 Awareness Political	5.64 (2.38)	4.62 (1.66) ^{j**}	6.29 (2.07) ^{j**}
8 Cynicism	4.75 (1.06)	4.54 (1.29)	4.43 (1.24)

Note. Pairwise comparisons for means with matched superscripts are significantly different (***) denotes significance at the $p < .001$ level, ** at the $p < .01$ level, and * at the $p < .05$ level)

A final exploratory analysis related to perceptions of compromise tested whether toxic and benign disinhibition would be related to perceptions of compromise. It was found that participants who perceived equal compromise had significantly lower levels of toxic disinhibition ($M = 2.07$, $SD = .99$) than those who perceived their partner as compromising most ($M = 2.57$, $SD = 1.89$), but not significantly lower than those who perceived themselves as having compromised the most ($M = 2.47$, $SD = 1.17$), $F(2, 179) = 3.189$, $p = .049$, $\eta^2 = .034$. No such relationship occurred according to levels of benign disinhibition $F(2, 178) = .713$, $p = .491$.

As the analyses for Research Questions 1a-g demonstrate, the conditions did not have an effect on the major outcome variables, so it was excluded from further modeling except as a control variable.

Hypotheses 1a-b and Research Question 2:

Hypothesis 1 predicted that (a) toxic, but not (b) benign, online disinhibition, would be related to performing less accurately on a questionnaire asking individuals to anticipate how their partner might have filled out an abortion attitudes questionnaire. This outcome variable, labeled “Other-Awareness,” was created using an index of correctly aligned responses between an individual’s pre-test Abortion Attitude Profile responses, and their partner’s guesses of how they might have responded. Correlation analyses were used to determine whether these variables were related.

H_{1a} predicted a negative correlation between toxic online disinhibition and Other-Awareness, which was supported. Regression analyses demonstrate that this relationship is not only correlational, but predictive in nature, albeit here having a very small effect; toxic online disinhibition explained 2.2% of the variance, $R^2 = 0.22$, $F(1,180)=4.12$, $p = .044$ —higher toxic disinhibition significantly predicted lower Other-Awareness ($\beta = -0.302$, $p < .05$). Also as predicted in H_{1b}, there was no correlation between benign online disinhibition and Other-Awareness, $r = .041$, $n = 182$, $p = .587$. Support for these two parts of Hypothesis 1 suggest, then, that the original subscales constructed by Udris (2014) can be used to identify circumstances under which negative conversational outcomes might occur during political deliberation. This finding lends support to the notion that when the online disinhibition process results in toxic disinhibition, it will ultimately also lead to decreased other-awareness, further inhibiting the deliberation process itself.

The second research question examined whether awareness of one’s online presence (SSAS - Online Presence, or simply SSASO) would be associated with either

type of online disinhibition. A correlation analysis found that SSASO was not correlated with benign online disinhibition $r = .030$, $n = 181$, $p = .685$, but it was (weakly) negatively correlated with toxic online disinhibition $r = -.144$, $n = 182$, $p = .049$.

Hypotheses 2a-b:

Hypothesis 2a-b predicted that men would have (a) higher toxic online disinhibition scores, and (b) be worse at guessing how their conversation partner would have filled out an abortion questionnaire than women. A MANOVA was conducted to test for main effects of gender on these two outcome variables. The multivariate result was significant for gender $F(2, 179) = 16.891$, $Wilks' \lambda = .841$, $p < .001$. Additionally, the univariate F tests showed significant difference between men and women for toxic online disinhibition, $F(1, 180) = 30.54$, $p = .001$, $\eta^2 = .145$ such that men had significantly higher toxic online disinhibition ($M = 2.73$, $SD = 1.16$) than women ($M = 1.89$, $SD = 0.86$); and men had significantly lower Other-Awareness ($M = 5.51$, $SD = 2.24$) than women had ($M = 6.28$, $SD = 2.02$), $F(1, 180) = 5.53$, $p = .020$, $\eta^2 = .030$. An exploratory moderation analysis was also run to test for an interaction between gender and toxic disinhibition on other-awareness, but the interaction was nonsignificant. Still, H_{2a-b} were supported.

Research Question 3

RQ₃ was analyzed using SPSS's "split file" function, comparing differences in stereotyping behaviors within Pro-Life and Pro-Choice individuals according to whether they were matched with an ingroup or outgroup member. A MANOVA revealed that within Pro-Life individuals there were differences on stereotyping in both matched-stance dyads and mixed-stance dyads, and that the prevalence of Pro-Life versus Pro-Choice

stereotypes were significantly different from one another across these categories, and while a similar difference occurred within Pro-Choice individuals, stereotyping outgroup members was more prevalent $F(2, 48) = 8.76$, $Wilks' \lambda = .733$, $p < .001$. As shown in Table 16, Pro-Life participants applied stereotypes relevant to their own ingroup (other Pro-Life individuals) when deliberating with an ingroup member more than they applied stereotypes relevant to outgroup members (Pro-Choice individuals) when deliberating with an outgroup member. The same was not true for, Pro-Choice participants, who instead applied relevant stereotypes to their outgroup members more than they applied relevant stereotypes to their ingroup members. Recall, stereotypes relevant to Pro-Life and Pro-Choice individuals (see Table 9) were created using a thought-listing task from the pilot study for this experiment. It should not be argued that these groups of stereotyping descriptors are perfect inverses of one another (for example, also some might argue that it is so, “feminist” and “sexist” are not considered equal opposites for the purposes of this study, since different groups produced these stereotypes and may perceive different connotations for these particular labels).

Table 16. Prevalence of stance-specific stereotypes used by participants

	Within Matched- Stance Dyads, M(SD)	Within Mixed- Stance Dyads, M(SD)
<i>Evaluator: Pro-Life Participants</i>		
Pro-Life Stereotypes**	4.06 (1.65)	2.31 (2.01)
Pro-Choice Stereotypes**	1.50 (1.51)	3.17 (2.23)
<i>Evaluator: Pro-Choice Participants</i>		
Pro-Life Stereotypes***	2.13 (1.90)	4.00 (2.55)
Pro-Choice Stereotypes*	2.86 (1.57)	2.19 (1.47)

Note: * Denotes between-subjects effects significant at the $p < .05$ level, ** at the $p < .01$ level, and *** at the $p < .001$ level

Hypotheses 3a-b:

Hypotheses 3a and 3b predicted that those higher in (a) toxic online disinhibition but (b) not benign online disinhibition would perceive their partners as more hostile. Correlation analyses found that toxic online disinhibition was indeed positively correlated with perceived hostility, $r = .192$, $n = 182$, $p = .009$; but benign online disinhibition was not related to perceptions of hostility, $r = .056$, $n = 182$, $p = .457$. Hypotheses 3a-b were therefore supported.

Hypotheses 4a-c:

Mediation analyses using Hayes's (2013) PROCESS macro for SPSS was used (all tests using 10,000 bootstrap samples, CI = 95%) to investigate the hypotheses that (a) perceived rationality, (b) perceived intelligence, and (c) perceived civility would mediate the relationship between toxic online disinhibition and perceptions of partner's hostility. Regression of toxic online disinhibition on the first hypothesized mediator, perceptions of partner's rationality, was significant, $b = -.19$, $t(179) = -2.26$, $p < .05$, $p = .040$ (LLCI = $-.3563$; ULCI = $-.0241$); the mediator (perception of partner's rationality), controlling for toxic online disinhibition, was also significant, $b = -0.63$, $t(179) = -8.28$, $p < .001$, $p = .000$ (LLCI = $-.7742$; ULCI = $-.4762$). Controlling for the mediator, toxic online disinhibition scores were not a significant predictor of perceived partner hostility, $b = .045$, $t(178) = .81$, $p = .4195$. A Sobel test was conducted and found full mediation in the model ($z = 2.17$, $p = .03$). It was found that perceptions of partner's rationality fully mediated the relationship between toxic online disinhibition and perceptions of hostility, thus H_{4a} was supported.

Regression of toxic online disinhibition on the second hypothesized mediator, perceptions of partner's intelligence, was not significant, $b = -.167$, $t(179) = -1.74$, $p = .084$ (LLCI = $-.3561$; ULCI = $.0228$). The mediator (perception of partner's intelligence), controlling for toxic online disinhibition, was significant, $b = -.443$, $t(178) = -7.42$, $p < .001$, $p = .000$ (LLCI = $-.5606$; ULCI = $-.3249$). Controlling for the mediator, toxic online disinhibition scores were not a significant predictor of perceived partner hostility, $b = .09$, $t(178) = 1.64$, $p = .1021$. Because only the b-path was significant, H_{4b} was unsupported.

Regression of toxic online disinhibition on the third hypothesized mediator, perceptions of partner's civility, was nearly, but not, significant, $b = -.128$, $t(179) = -1.90$, $p = .059$ (LLCI = $-.2604$; ULCI = $.0049$). The mediator (perception of partner's civility), controlling for toxic online disinhibition, was significant, $b = -.734$, $t(178) = -9.58$, $p < .001$, $p = .000$ (LLCI = $-.8847$; ULCI = $-.5824$). Controlling for the mediator, toxic online disinhibition scores were not a significant predictor of perceived partner hostility, $b = .07$, $t(178) = 1.29$, $p = .1992$. Because again only the b-path was significant, H_{4c} was unsupported.

Within H_4 , only perceptions of a partner's rationality fully mediated the relationship between toxic online disinhibition and perceptions of hostility, such that as toxic online disinhibition scores increased, participants perceived their partners as less rational, resulting in higher perceptions of hostility. The same was not true of perceptions of intelligence or civility, although both lower perceptions of intelligence and lower perceptions of civility did significantly predict higher perceptions of hostility.

Hypotheses 5a-i and Research Question 4:

Hypotheses 5_{a-h} posited that participants matched with a partner of opposing abortion stance would perceive their partners as (a) less rational, (b) less intelligent, (c) less civil, and they would (d) feel colder, (e) identify with less, and come away with more feelings of (f) perceived hostility, (g) political deliberation cynicism, achieve (h) less other-awareness and finally, (i) less task success. In order to test for main effects of dyad composition on these variables, a MANOVA was performed. Results of the MANOVA indicated that there was a significant effect of dyad composition on partner evaluations and other outcomes, $F(16, 334) = 2.15$, $Wilks' \lambda = .775$, $p < .001$. Significant differences occurred according to dyad composition on all of the variables examined in H₅, in the direction predicted, except for political deliberation cynicism. Pairwise analyses for this MANOVA can be found in Table 17. Notably, there were no significant differences between dyads composed of both Pro-Life individuals and both Pro-Choice individuals. All differences occurred between either both Pro-Life and Mixed, or Pro-Choice and Mixed dyad compositions, providing evidence for support of H₅, which predicted dyads comprised of ingroups (both Pro-Life and both Pro-Choice) would hold more favorable perceptions than dyads comprised of outgroups (Mixed dyad composition).

For Hypothesis 5_i, a chi-square analysis was used to determine whether dyads composed of opponents (outgroup dyads) would have more success completing the deliberation task than dyads composed of matched-stance partners (ingroup dyads). Recall, “success” was determined by whether both members of each dyad, upon completion of the task, were able to independently choose the same policy options—the ones they had agreed upon during the deliberation period. The chi-square analysis

revealed, interestingly, that dyad composition did not have an effect on whether or not task success occurred, $\chi^2(1, N = 182) = 0.403, p = .525$. Thus, H_{5i} was unsupported.

Because some significant differences were found in the sub-hypotheses of H_3 , exploratory correlation analyses were run to test whether more polarized individuals would be predisposed to specific outcomes. Polarization was measured by taking the eight-point abortion stance scale (one being fully “Pro-Choice,” and eight being fully “Pro-Life”), and assigning new values to this scale such that the outer two, most extreme options became fours, the next two inner options became threes, the next two inner options became twos, and the center two options became ones, thus creating a four-point scale in which one referred to the most centrist stances, and four referred to the most polarized stances. However, this “polarization” scale did not correlate with any outcome variables, save one: positively-valenced emotions. That is, those with more polarized stances, regardless of which stance was held or whether the participant was assigned to debate with an outgroup member, generally came away from the task with more positive emotions, perhaps suggesting that these participants were highly invested, or glad to debate $r = .182, n = 178, p = .015$. There was, however, no effect of polarization on perceptions of civility, rationality, intelligence, or hostility, or on political deliberation cynicism, identification with one’s partner, other-awareness, task success, or “trait” variables such as toxic or benign online disinhibition, self-awareness of social groups, self-awareness of online presence, or negatively-valenced emotions following the task.

Table 17: Results of Hypotheses 5a-h MANOVA

<i>H5 section</i>	Individuals Within Dyads of Differently Composed Abortion Stances		
	Both Choice (<i>n</i> = 91)	Both Life (<i>n</i> = 15)	Mixed (<i>n</i> = 70)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
a Perception of Partner's Rationality	6.44 ^{a***} (.65)	6.36 (.56)	5.76 ^{a***} (1.12)
b Perception of Partner's Intelligence	5.98 (.97)	5.97 ^{b***} (1.15)	5.23 ^{b***} (1.13)
c Perception of Partner's Civility	6.47 ^{c***} (.60)	6.38 (.61)	5.5.9 ^{c***} (.95)
d Thermometer Toward Partner	84.56 ^{d***} (13.29)	87.4 ^{e*} (12.97)	73.44 ^{d***e*} (21.74)
e Identification with Partner	5.87 ^{f***} (.81)	5.91 ^{g*} (.70)	5.29 ^{f***g*} (1.01)
f Perception of Partner's Hostility Political	1.52 ^{h***} (.69)	1.58 (.82)	2.07 ^{h***} (1.09)
g Deliberation Cynicism	4.37 (1.15)	4.64 (1.47)	4.61 (1.26)
h Other-Awareness	6.62 (1.84) ^{i***}	6.77 (2.05) ^{j*}	5.16 (2.29) ^{i***j*}

Note. Pairwise comparisons for means with matched superscripts are significantly different (***) denotes significance at the $p < .001$ level, and * at the $p < .05$ level)

Research Question 4 asked whether there would be an interaction between dyad composition and abortion stance on these variables. The MANOVA used for analysis of H_5 was performed once again, this time using the split file function of SPSS, additionally comparing groups according to individuals' abortion stance. Results of this MANOVA indicated that there was a significant effect of dyad composition on partner evaluations and other outcomes within Pro-Choice individuals, $F(7, 118) = 2.45$, *Wilks' λ* = .873, $p < .05$ ($p = .037$); but there was not enough power to detect differences within Pro-Life

individuals $F(7, 42) = .91$, $Wilks' \lambda = .868$, $p = .509$, observed power = .342. Within Pro-Choice individuals, participants in matched-stance dyads perceived their partners to be significantly more rational ($M = 6.44$, $SD = 0.67$) than individuals in mixed-stance dyads perceived their [Pro-Life] partners to be ($M = 5.78$, $SD = 1.16$), $p < .001$ ($p = .000$); more intelligent ($M = 5.97$, $SD = 0.96$) than individuals in mixed-stance dyads perceived their [Pro-Life] partners to be ($M = 5.24$, $SD = 1.39$), $p < .001$ ($p = .000$); more civil ($M = 6.47$, $SD = 0.65$) than individuals in mixed-stance dyads perceived their [Pro-Life] partners to be ($M = 5.89$, $SD = 1.02$), $p < .001$ ($p = .000$); had warmer feelings toward their matched-stance partners ($M = 84.56$, $SD = 1.74$) than individuals in mixed-stance dyads had toward their [Pro-Life] partners to be ($M = 73.43$, $SD = 2.80$), $p < .001$ ($p = .000$); felt higher identification with their matched-stance partners ($M = 5.87$, $SD = 0.09$) than individuals in mixed-stance dyads had with their [Pro-Life] partners to be ($M = 5.29$, $SD = .15$), $p < .001$ ($p = .000$); and finally, perceived less hostility from their matched-stance partners ($M = 1.52$, $SD = 0.09$) than individuals in mixed-stance dyads perceived from their [Pro-Life] partners to be ($M = 2.07$, $SD = .14$), $p < .001$ ($p = .000$). Dyad composition did not have an effect on political deliberation cynicism within Pro-Choice individuals, however $F(1, 124) = 0.013$, $p = .909$.

Research Question 5a-b:

Research Question 5a asked whether self-awareness of one's own social groupings (SSAS – Social Identity, or simply SSASSI) would be related task success. An ANOVA was used to determine whether this relationship existed, but there was no significant difference between those who were successful (of which there were 76 cases, 38 dyads) or unsuccessful (106 cases, 53 dyads) in their awareness of social identity $F(1,$

179) = .101, $p = .751$, observed power = .062. Research Question 5b asked whether polarization would have any effect on task success (i.e., participants in dyads successfully able to come away from the task and, individually, each report that the same set of policies had been agreed upon during the task were considered to have achieved task success). An ANOVA revealed that there was a significant difference between successful and unsuccessful participants, such that successful participants slightly more polarized attitudes ($M = 2.83$, $SD = 1.17$) than unsuccessful participants held ($M = 2.45$, $SD = 1.09$), $F(1, 180) = 4.96$, $p < .05$ ($p = .027$), $\eta^2 = .03$.

Research Question 6:

Research Question 6 asked whether attitude polarization would be related to higher awareness of one's social identity. A correlation analysis found that there was no relationship between polarization and SSASSI, $r = .041$, $n = 181$, $p = .586$.

CHAPTER 5.

DISCUSSION

The results of this research yielded a great many connections between participant characteristics, emotional responses following ingroup/outgroup online contact, and perceptions of such experiences. Together, these connections provide evidence that individuals with higher levels of toxic online disinhibition are more likely to perceive online interactions and conversation partners less favorably than those with lower levels of toxic disinhibition. However, it was also found here that such individuals will respond well (that is, with less negative, defensive behaviors and perceptions) to partners who behave rationally—there including, nonuse of stereotypes and nonlinear argument patterns—and who do not over-compromise, but instead commit to standing their own ground in a debate. The particulars of such interactions, which should be taken as evidence of civility and productivity that can exist in anonymous, deindividuated, highly polarized settings, follows.

This study examined a number of theoretical frameworks not against one another, but through one another. Beginning with Suler's (2004) online disinhibition, it was expected that those with higher predisposition to beliefs consistent with toxic disinhibition would behave and perceive the task and their partner differently. It should be noted that one's predisposition to toxic disinhibition was measured on a pre-test scale, days before each participant engaged in the actual task, and therefore should not be considered a state prompted by the stress of the task, outgroup contact, or heightened awareness due to lab settings—though there is no way here to control for the online behaviors each person engaged in that day. So, these disinhibition scales—toxic and

benign—are discussed as though they measure trait disinhibition, with caution and acknowledgment that these levels might change somewhat depending on digital technology use and changing circumstances.

With that in mind, Hypotheses 1_{a-b} found that higher toxic, but not benign, disinhibition predicted individuals performing worse on an other-awareness test. Recall, “other-awareness” was measured by instructing participants to fill out the abortion attitude profile questionnaire as they assume their partner would have filled it out, and then scoring this against the pre-test abortion attitude profile of their dyad partner (i.e., tallying how many answers were in alignment). This finding suggests that toxic disinhibition beliefs include those that might preclude an individual from understanding a debate partner’s argument, or perhaps from even really listening to (here, reading) it. Solipsistic introjection, one of the six characteristics of the disinhibited state described by Suler (2004) may be responsible for this effect, as imagining one’s conversation partner by transferring internally-generated assumptions and projecting traits of others onto an anonymous partner may result in the glossing over of actual arguments made by said partner. An attempt was made to create a scale isolating this particular construct, and another scale to differentiate between the imagined *person* and the assumed *argument*, but these scales did not behave as expected, therefore future work may still be needed to examine these nuances.

Research Questions 1_{a-g} found that the conditions created in this study did not have an effect on any outcome variables, except some incredibly small evidence that Pro-Choice individuals are predisposed (i.e., the finding occurred in the control condition only) to polarizing when talking to someone of the same stance (although this finding

existed in a post-hoc analysis for which the omnibus test was not significant, so this “evidence” is weak and should be re-assessed in a more potent experimental design). The conditions were meant to prime dyads to engage in the same task with different sub-goals—either to understand their partners (the “understanding” condition), to get their partners to understand them (the “demonstrative” condition), or were given no additional instructions beyond the task (the control). It cannot be said that these frames have no effect—it may easily be that the instructions distinguishing each condition were not strong enough to elicit a behavioral response. Indeed, only one line—bolded—was added to the task instructions in each treatment condition. The task itself may have been difficult or stressful enough that incorporating these additional goals during a fifteen minute time period resulted in cognitive overload and was discarded. Alternatively—and this is unfortunately common, it seems, in experimental designs that rely on text—people did not read the prompt as carefully as they were meant to.

Perceptions of compromise were strongly linked with many perceptions of one’s conversation partner, even when controlling for the stance-composition of each dyad. As was seen in Table 15, participants had more favorable perceptions of their partner and of the experience when the compromise was seen as relatively equal. Interestingly, participants appeared to prefer conversations in which equal compromise occurred more than they preferred compromise that would serve their benefit—the compromise of their partners—suggesting both that individuals are sensitive to equality in a conversation’s outcomes, and that they do not want to feel as though they have dominated a conversation. One possible reason for this finding could be that debate with a partner who

seems like too much of a “pushover” makes many people feel as though their engagement has not been met with as much interest.

Beyond categorizations related to Pro-Life and Pro-Choice identities, political affiliation, gender, or any other immediately distinguishable social identities measured here, it almost seems as though an additional social identity delineation becomes evident: more engaged citizens and more apathetic citizens. As was concluded by Neblo and colleagues (2010), there may be a specific type of social identity related to people’s willingness to engage in deliberative participation (see also: Mutz & Martin, 2001), and often these are the same people who rebuke partisan politics. This picture becomes more evident in consideration of the specific differences found on outcome variables according to perceptions of compromise: People who wanted to engage (who often had higher toxic disinhibition) did not enjoy a partner who did not appear to want to engage, even if their partner’s disengagement allowed them to “win” the discussion (i.e., compromise less). That is, their desire for sparring was not satisfied, even if they held polarizing views that may have been threatened if sparring *had* occurred. This is interesting, as higher polarization may seem as though it would go hand-in-hand with higher toxic disinhibition—here, it is shown that “winning” an argument does not simultaneously satisfy the desires produced by polarization and toxic disinhibition.

Those who, in response to the question “Who do you feel had to compromise most during the task?” answered, “Neither – we compromised equally” gave more favorable views across the board, with the exception of political cynicism, again for which there was no significant difference according to perceptions of compromise. Those who felt they had compromised equally gave higher evaluations of their partner’s civility,

intelligence, rationality, perceived their partners as less hostile, identified with their partners more, had warmer feelings toward their partners, and had better other-awareness scores, suggesting that mutual compromise also made it easier to understand an opponent's views. On all but one (other-awareness) of the variables with significant differences according to perceptions of compromise, a significant difference occurred between people who answered "neither" and people who answered "me." This makes immediate sense, considering that equality would be most desirable, and given our bias to notice inequalities that affect us over others, people would have the least favorable views of partners who made them feel that they themselves had to compromise. Some variables—perceptions of rationality, feelings of warmth, and identification with partner—had significant differences between those who answered "me" and "neither," as well as between "my partner" and "neither," but there were no variables with significant differences between those who answered "me" and "my partner." In general, the trend was that equality yielded the most favorable results, followed by a partner's compromise, and finally one's own compromise. Only on one variable—other-awareness—was this trend broken, with only a significant difference between those who answered "neither" and those who answered "my partner," this time with those who answered "me" having middling other-awareness scores.

As mentioned above, it may be the case that people did not like overly-compromising partners as much as they liked equally-compromising partners because overly-compromising partners may appear to be less engaged or less genuine. The fact that participants with overly-compromising partners also were less able to accurately fill out the abortion attitude profile as they believe their partner would have filled it out lends

evidence to this conjecture. If a non-resistant participant gave very little for their partner to respond to, it would make sense that they would appear to be the more compromising one, and give their partner too little insight into their own beliefs for them to adequately guess their specific attitudes—thus, lower other-awareness scores for the partners of more compromising participants. Alternatively—and this feels unlikely—many people who are truly worse at listening to or understanding an online debate partner’s beliefs are also self-aware enough to recognize that they dominated the conversation.

The second research question found that higher awareness of one’s online presence—or, given the self-report nature of this awareness, perhaps higher confidence in one’s own awareness of one’s online presence—was negatively correlated with toxic online disinhibition, suggesting that those who are more cognizant of their behavior online are less likely to hold beliefs consistent with toxic disinhibition. The non-relationship between SSASO about benign disinhibition suggests that this higher awareness of one’s own online behavior does not impede the disinhibited state entirely—only, it makes toxic beliefs and behaviors less likely. Higher SSASO scores may, as mentioned, actually be referring to higher confidence in one’s own online self-awareness, or alternatively may indicate a higher belief that online self-awareness is socially desirable. In either case, the way that this variable associates with lowered toxic disinhibition scores but undisturbed benign disinhibition scores suggests that higher introspection on one’s online habits, even if occurring for the sake of appearing more socially desirable, may have positive effects on online deliberation. More outcomes relating specifically to toxic online disinhibition will be discussed shortly, but for now I

turn to an examination of toxic disinhibition as it occurred along gendered lines in the sample.

Hypotheses 2_{a-b} found that men had higher toxic (but not benign) online disinhibition scores, and lower other-awareness scores than women. That is, men willingly admitted to engaging in more hostile or assumptive behaviors in online settings than women admitted to, which may reflect the communication socialization messages given to young women versus young men—that women should listen and hedge, and men should be assertive—or it may be reflective of women being less willing to admit to toxic behaviors online, even though it is known that women and young girls perpetrate social aggression far more often than men and young boys do, and far more than they perpetrate physical aggression (Crick, Bigbee, & Howes, 1996; Crick & Grotpeter, 1995; Martins & Wilson, 2012). Nonetheless, it appears that toxic disinhibition and its relationship to reduced other-awareness especially affected men. As discussed above, it was found that toxic disinhibition, comprised of some beliefs supportive of explicitly hostile behavior, and some supportive of assumptive behaviors, ironically yields less other-awareness in the very same people who believe that assumptive behaviors are “okay” to engage in. Perhaps it is that these types of people do not believe their other-awareness will be affected, or perhaps other-awareness is not the metric by which they themselves believe a successful outcome has occurred.

In addition to gender differences found in this study, other social identities may have affected outcome variables as well. Because Pro-Life and Pro-Choice identities have ill-behaved, ectoplasmic borders at best, existing in an ever-malleable state at the intersection of gender, religiosity, political party, age, and several other identities, it was

first decided to examine these groups more broadly, as ingroups and outgroups. To skip ahead briefly in the order of hypotheses, the fifth set found that participants within dyads of mixed abortion stances had worse evaluations of their partners' civility, intelligence, and rationality, perceived more hostility from their partners, felt colder toward them, identified with them less—and yet, these effects did not produce stronger online political deliberation cynicism. This effect occurred regardless of pre-test abortion stance differences, and this effect occurred independently from the influence of pre-existing beliefs consistent with higher proclivity toward toxic disinhibition. This was not an unexpected finding, as it upholds findings consistent with previous social identity research—that individuals will give preference to members of their ingroup rather than their outgroup. The task performed in this experiment would naturally be easier to accomplish with a politically similar individual, as it would require less compromise or defense of one's own stance, so a distinction on these variables by ingroup/outgroup dyad composition is theoretically sound. However, the fact that unfavorable perceptions did not result in higher political cynicism is interesting both in that the negatively-valenced scales used in this study overwhelmingly tended to track together, whereas online political deliberation cynicism, which was deemed a statistically sound scale, did not. This could be good news for those who believe more outgroup contact, online debate, or really any politically charged conversation online regardless of whether it occurs with outgroup members, would produce more exhaustion and avoidance of online engagement. Now, this cynicism was also not alleviated—it just remained unaffected—so it cannot be argued that online deliberation whets a thirst for more of the same.

Longitudinal analyses in this area would be desired, to test whether extended, repeated exposure to such online engagement has an effect on cynicism.

The third research question found that Pro-Choice individuals stereotyped Pro-Life individuals more than Pro-Life individuals stereotyped Pro-Choice individuals, but there was no resulting difference in other-awareness. In fact, Pro-Life individuals also more heavily stereotyped members of their own ingroup than they stereotyped members of their outgroup. In the context of H₄, this could be troubling if these stereotyped beliefs are related to stereotype use during debate itself, which those who have higher levels of toxic disinhibition might view as irrational behavior. There are a number of reasons why Pro-Life individuals may have been more heavily stereotyped by their partners. For one, on most left-leaning college campuses, Pro-Life students will be in the minority (Ertelt, 2011), and minorities are commonly stereotyped. Echoing the assertions of Fiorina and colleagues (2005), it was found in the pilot and full study samples for this project that Pro-Life individuals comprised a narrower range of social identities and policy preferences, whereas Pro-Choice individuals comprised a much more scattered group of identities, and included a much more varied range of policy preferences (although indeed, there was some overlap and even some cases of the most lenient self-described Pro-Life individuals allowing more exceptions to an abortion ban than the most non-lenient of self-described Pro-Choice individuals were willing to allow, suggesting that these labels do not perfectly describe one's exact policy preferences). As previous scholars have illustrated, the "Pro-Choice" identity is especially complex, as many of those who self-identify this way still "believe that abortion is wrong" but "nevertheless support the principle of a woman's right to choose as embodied in Roe" (Fiorina, Abrams, & Pope,

2011, p. 81). On the other hand, as was also found in this study, self-identified Pro-Life individuals were generally consistent in their support of restrictive abortion policy.

However much Pro-Choice individuals believed stereotypes applied to their Pro-Life counterparts, relying on these cognitive shortcuts did not appear to affect Pro-Choice participants' ability to guess how their partners would have filled out an abortion attitude profile (their "other-awareness"). Pro-Choice individuals appear to be stereotyping their outgroup members' personas, but not the arguments of their outgroup members, which again in the context of H_{4a-c} , may bode well for such outgroup contact as long as stereotypical descriptors are not brought into the debate itself to put off Pro-Life individuals with higher toxic disinhibition.

Research Question 4 lends nuance to the findings of H_{5a-i} . There was an interaction between dyad composition (matched-stance/ingroup, mixed-stance/outgroup) and abortion stance such that the effects found above were primarily seen within Pro-Choice individuals. Pro-Choice participants saw matched-stance partners as more civil, more rational, more intelligent, felt more warmly toward them, identified with them more, and perceived less hostility; however, there was still no effect on cynicism, suggesting at least that here, these perceptions still do not affect cynicism. The same effects did not occur in Pro-Life individuals in matched-stance dyads, nor did they occur in Pro-Life or Pro-Choice individuals who were in mixed-stance dyads. It may be that, given the relatively smaller number of Pro-Life individuals in the sample in general, there was simply not enough power to detect any differences. As effect sizes were small to middling even for the larger Pro-Choice sample, this could easily be the case. Alternatively, the aforementioned effects of political climate as they might uniquely

affect each group could have caused Pro-Choice individuals to feel relief, and thus a more pronounced camaraderie with matched-stance partners.

Research Questions 5_{a-b} and 6 demonstrated that strong social identity awareness and polarization work in independent ways, and should not be conflated. Having a strong sense of one's social identity, and groups that comprise one's social identity, did not have an impact on task success even when controlling for mixed-stance (outgroup) dyads, suggesting, as might be expected, that task success is not driven by conflicting identities—but also, it is not impeded by it. Dyads with more highly polarized individuals, even when controlling for outgroup status—in which case the delta between attitudes would be quite large—were more successful in completing the task than dyads with individuals whose preferences were more centrist. This may be because, as previous research has suggested, polarization may function to increase political interest (see: Abramowitz, 2010; Mutz & Martin, 2001), which then emboldens individuals to lean in to higher-stakes deliberation with a commitment to productive outcomes.

Hypotheses 3_{a-b} provides further validation that the online disinhibition scale developed by Udris (2012) is sensitive enough to delineate beliefs that could produce a toxic versus merely benign disinhibition effect, as distinguished by Suler (2004). As hypothesized, those with higher levels of toxic disinhibition perceived more hostility from their partners, yet higher levels of benign disinhibition did not. Additional analyses found that pre-test toxic disinhibition (again, which may be considered *trait* toxic disinhibition since it was not prompted by the treatment—the caveat being that it should generally be considered a state) although behaved much the same as outgroup contact, and benign disinhibition was not correlated with any one of these variables. That is,

higher levels of toxic disinhibition were correlated with the very same variables as found in H_{3a-g}, including a non-significant correlation with political cynicism. What this finding may suggest is a need for more research that examines the cyclical relationship between extended online outgroup contact, and higher levels of toxic disinhibition.

Interestingly, though, the non-correlation found between toxic disinhibition and online political deliberation cynicism suggests that, if indeed such a cycle of expecting and subsequently perceiving more hostility exists, it is not related to online political deliberation cynicism. Again, this may be good news for those who worry that extended political argument online will produce such cynicism—but it may be bad news for those who worry that it will produce more hostile perceptions. The divorcing of these two concepts—cynicism about online political deliberation and toxic disinhibition—is interesting, as it suggests individuals can be cynical without engaging in toxic behaviors online, if only such cynical individuals can be led to engage. Conversely, one can behave in toxic ways without being a cynic, therefore perhaps behavioral changes could be encouraged without taking away one's desire to engage. Presently, very little research exists that examines online political deliberation cynicism—political cynicism scholarship generally focuses on that which is aimed at politicians (Cappella & Jamieson, 1997; de Vreese & Semetko, 2002), political parties (Dalton & Wattenberg, 2000), and media outlets (Cappella & Jamieson, 1997; Moy & Pfau, 2000; Valentino, Buckmann, & Buhr, 2001). It has been mused upon by scholars, that positive outgroup interactions—even those containing disagreements—might decrease such cynicism (Eveland, Morey, & Hutchens, 2011). This experiment, however, was unable to find evidence for such a hypothesis, although the political deliberation cynicism scale used here may provide

future direction for researchers attempting to understand this particular subset of political cynicism.

Although there is no “magic intervention” that can allow us to fully reap the benefits of benign disinhibition without occasionally sparking the conditions for the toxic counterpart of this process, thus far it appears that cynicism and polarization—although commonly described as ingredients for hostility—will not produce toxic behaviors, especially in individuals who have higher online self-awareness.

Hypotheses 4_{a-c} give additional context to the mechanisms through which a predisposition to toxic disinhibition produces less favorable perceptions of a conversation partner (regardless of ingroup or outgroup status). Here it was found that perceptions of a partner’s rationality fully mediated the relationship between toxic disinhibition and perceived hostility, whereas perceptions of intelligence and civility did not seem to act as mediators toward this outcome. The definition of deliberation used for the purposes of this study, in an attempt to keep with a Habermasian approach to rational-critical debate, requires that both rationality and civility be integral parts of any such conversation for it to be considered productive.

Should it then be concluded that individuals more predisposed to beliefs consistent with toxic disinhibition will always be at a disadvantage in their attempts to deliberate online, because this state of mind is so strongly connected to a decrease in perceptions of a partner’s rationality and civility (and intelligence)? Perhaps this is so, yet their inclination to see their debate partners as less civil did *not* mediate a path toward perceptions of hostility, while their inclination to see their debate partners as less rational did. This is especially interesting, considering civility is often considered the diametric

opposite of hostility, so it was this construct in particular that was expected to act as the strongest mediator of the three. However, it may make theoretical sense that those who espouse toxic beliefs about online communication themselves are consistent in how they perceive the behaviors of a partner—that is, they make allowances for a certain level of incivility because they don't expect or require civility to occur in the first place. It has also been argued that a third-person effect exists with regard to civility—that is, that people generally believe demonstrations of civility can be persuasive to others, but that it is not persuasive or important to oneself (Chen & Ng, 2016). What individuals with high toxic disinhibition cannot tolerate, however, is irrationality in a debate—they perceive this breach in logical flow as hostile. Still, perceptions of intelligence also did not mediate this effect, and so the reason individuals with higher toxic disinhibition see irrationality as hostile cannot be tied to seeing their partner as stupid or ill-equipped to debate. It may then be that such people believe irrational partners are behaving irrationally on purpose, making jumps in logic or inconsistencies in reason purposefully, to skew the debate in their favor.

So, perceiving one's partner as uncivil may not be a damning factor in online conversations had between those with higher levels of toxic disinhibition, if perceptions of hostility are to be considered any indicator of how such people decide whether a deliberation partner has “played fair,” so to speak. This may appear to be inconsistent with the original Habermasian approach to deliberation, yet when one considers that—regardless of this particular mediation—perceptions of civility, rationality, and intelligence were strongly correlated, even for people with high levels of toxic disinhibition, it follows, then, that for these people in particular, rationality *is* civility.

Indeed, scholars have described “highly rational” debate as that which provides “detailed point-by-point rebuttals of quoted sentences and paragraphs,” suggesting that a rational debater recognizes and respond to another’s words—doubling as an act of civility (Kelly, Fisher, & Smith, 2009).

If future research finds that toxic disinhibition is cyclically produced in people who have come to expect toxic behaviors from others by repeated engagement with those who do, an explanatory factor—a corner piece to the puzzle that is the intersection between online disinhibition (Suler, 2004) and deliberation theories—may then be to examine social media users’ expectations in debate (Deegan, Hehman, Gaetner, & Dovidio, 2015; Hwang, Kim, & Huh, 2014), and whether irrational behaviors such as nonlinear arguments and logical fallacies are the likeliest to trigger toxic responses from others.

Theoretical and Practical Implications

Overall, the findings presented in this study demonstrate that there are strong personality differences, in addition to those related to social identity, that affect the ways toxic disinhibition will change one’s perceptions of others online, and their feelings about their own online behavior. For one, it was found that the very same people who believe toxic behaviors online are more acceptable and less damaging are the most likely to have lessened post-debate accuracy of understanding their partner. It was also found that irrational behaviors are deemed indicators of hostility for such people, whereas incivility and unintelligence are not, suggesting that the very effect that bothers them the most—illogical statements, perhaps borne of assumption, stereotyping, and misunderstandings—are likely to be perpetrated by they themselves. This process creates a cycle where

demonstrations of toxic disinhibition generate more perceptions of hostility, and thus perhaps reinforced beliefs consistent with toxic disinhibition, regardless of whether past or future online interactions have or will occur with outgroup members.

Additionally, these results should be taken in the context of late 2017, still near the beginning of the Donald J. Trump administration, when Democrats and left-leaning individuals were likely to feel a dip in confidence and political efficacy, perhaps causing a more guarded approach to conversations with outgroup members. On the other hand, Pro-Life individuals may have had unique reasons for the experimental topic itself—abortion—to cause defensive behaviors or negative emotions, as abortion has been (for the most part) legal for as long as any of the participants in this study would have been alive. It could be expected that if any of these circumstances were different—a different administration, different Federal abortion laws, or perhaps even an entirely different topic—the connections made in this study between online communication processes and relevant social identities might have been different. Recall, abortion was chosen as the topic through which to put these media effects theories in concert with one another, simply because it was found that this topic was most divisive—so while these findings may be used as a template for understanding the intersections of the theories themselves, no assertion will be made that these states of disinhibition, deindividuation, or self-awareness are inherent traits of any social identities easily affected by the political climate. This study is meant to give insight into online deliberation processes alone, with implications for online ingroup and outgroup contact more broadly stated.

It was concluded here that just as benign and toxic disinhibition will not always occur in the same spaces (echoing Udris, 2011), so too are toxic disinhibition and attitude

polarization not to be conflated. To wit, benign disinhibition can occur in polarized individuals (who are more prone to deindividuation effects, see Postmes, Spears, & Lea, 1998), discussing polarizing topics, without toxicity. As mentioned, previous literature in the realm of online disinhibition and deindividuation has predominantly concluded that civility occurs more often in spaces that are not deindividuated (that is, between people whose personal identities are not superseded by their social identities) and that are not supportive of online disinhibition (that is, less anonymous, invisible, and so on, and which encourages less assumption). This study, however, aimed to illustrate the conditions of civility that *does* occur in these disinhibiting, deindividuated settings despite being theoretically improbable, or at least less common. The predominant contributions of this study suggests that the most potent factors promoting civility and thus, deliberative qualities of political communication online, can be positively influenced on the conditions that (1) interactants exhibit situational self-awareness of one's online behaviors, (2) interactants commit to engaging in rational behavior, inherent in this being an intent to avoid stereotype use, and (3) interactants are equally committed to cooperation and, perhaps ironically, do not over-compromise. In sum, it may appear that adherence to the old adage that it is preferable to "agree to disagree" accomplishes the exact opposite of its intent, at least for many people—instead of demonstrating respect for one another by recognizing one another's arguments and rising to meeting them rationally, with commitment, and in full awareness of the risks of online hostility, agreeing to disagree may only reinforce existing rifts between ingroups and outgroups.

Limitations and Directions for Future Research

This study was limited, in part, by a subtle manipulation, and therefore the extent to which it can be said that these debate frames do or do not have an effect on online intergroup contact outcomes is yet unknown. The subtlety of this manipulation was meant to increase external validity by mimicking the type of light encouragement one might encounter in the “rules” of an online forum, chat room, or other digital space that nudges interactants toward more or less understanding of one’s partner. As such, it may very well be that such encouragement, when too gently applied, do not have the outcomes desired by forum or chat moderators. Future designs might consider using much more potent primes to prod participants to interact in the desired ways, although more potency here would then perhaps test these debate frames at the expense of external validity. Ongoing analysis of chat logs from this study will help determine exactly what, if any, of the instructions from either treatment condition resulted in variance of debate content, which may then give some insight into when one-time instructions do or do not have an effect. One indicator of the conditions’ potency yet unexamined may be the extent to which participants use a discussion of criteria, rather than policy, to meet on common ground rooted in shared intentions. Criteria-based discussion may allow interactants to speak productively about a controversial topic without an end goal of compromise, which in many cases does not resemble a truly centric position (Kroll, 2005).

An original iteration of this study’s design was to manipulate whether dyads would be explicitly instructed to discuss policy outcomes or criteria for such outcomes, but it was ultimately decided that such a distinction may be better suited to longer-form manipulations, such as those tested for the use of semester-long curricula. Reframing the

deliberative process as a means of establishing criteria, rather than attempting to discuss solutions before criteria, may even be useful practice in teaching children foundational scientific principles. Between formulation of hypotheses and the actual testing of hypotheses, primary education does not as commonly delve into discussions of criteria application. Such additions to early education may help foster healthy deliberative habits on and offline. In both science and debate, outcomes may depend entirely on the formation of evaluative criteria and operationalization; skipping instruction related to this part of the scientific method may have detrimental effects on the public's capacity to deliberate effectively, causing frustration, distress, and perhaps even the polarization of those with differing attitudes. Thus, experimental research linking early science education to later online argumentation habits could be a fascinating place to begin piecing together the conditions predictive of later toxic online disinhibition beliefs.

Additionally, this study was limited to a Midwestern college student sample, though it would be interesting to examine a similar procedure for older or younger populations. Manipulation of various other technological characteristics related to disinhibition, such as anonymity, invisibility, and even varying levels of asynchronicity when paired with long-term use would also be useful. Additionally, manipulations of social support, such as the addition of “likes” and “dislikes” would aid in our understanding of the disinhibition effect as it interacts with the deindividuation effect. Finally, an examination of what prompts uncertainty reduction behaviors such as self-disclosure and question asking—especially in highly political, anonymous, deindividuated online debate, where merely one-sided demonstrative behaviors may be more immediately gratifying—would complement a similar design and help create a

clearer image of what exactly *benign* disinhibition can accomplish—indeed so little research in our field focuses on prosocial outcomes occurring in politically tense deliberation. Perhaps so far we have not the faith to expect that they could.

Future work should build upon this study by assessing what circumstances, experiences, and individual differences beyond those collected here would lead one to become predisposed to beliefs consistent with higher predisposition toward toxic online disinhibition, which, as shown, has reverberating effects on—if not the actual content of an online confrontation—the perceptions one comes away with immediately following deliberation in anonymous digital spaces. Beyond socialized identities such as gender, or context and era-dependent states such as political affiliation or identity with a policy, what might we expect occurs during one’s formative years, before such identities have crystalized, that would give one the inclination to espouse beliefs about online communication that would put them more at-risk for toxic disinhibition? Research in this area should now more strongly focus on the habits and parenting styles of households with young children and adolescents.

Concluding Remarks

The theories that drive media psychology centered on digital communication must take into account *both* the static features of technology and the ever-changing mental states of the user, and yet scholars who search for a solution to the problem of toxic online behaviors may often be tempted to consider one of two paths—encourage cultural change or, perhaps more drastically, change our technology. If indeed anonymity is one of the key predictors of online disinhibition, the latter path would suggest that digital media users made mandatorily identifiable would fix the problem (or at least prevent it).

Aside from the fact that such a mandate would likely never get much support due to the priority we place on our privacy, the findings of this study suggest that benign disinhibition, even in anonymous and heated debate, can be a neutral or even positive characteristic—and as was found here, it is not guaranteed that toxic disinhibition certainly follows from benign disinhibition. What this suggests is that it may not be a fool's errand to seek ways to enjoy the benefits of a disinhibited state while discouraging toxic disinhibition. However, in order to preempt additional hostility that may occur once oppositional social identities have formed, discouragement of beliefs associated with toxic disinhibition must begin young.

Still, it may not be enough to tell young social media users to “play nice”—after all, people are reluctant to believe that their beliefs and behaviors are influenced by media (Davison, 1983). My own ongoing work in the area of Third Person Effect—which, as in H₂ of this study, similarly finds that men (and conservatives, who may see themselves as particularly guarded against their perceived “liberal media bias”) are more affected by media use but less willing to believe that they are—will continue to seek solutions to the practical concerns of media use and exposure that leads to hostile online outgroup contact.

Conflict resolution in the first place, let alone online conflict resolution, is hardly discussed in the classrooms of K-12 schools. Perhaps it ought to be. Perhaps we need not wait for the precedent of mandatory debate or ethics to be first introduced in younger classrooms—it may be enough that students already learn to develop proofs, write essays, and conduct experiments, as are skills learned in traditional math, language, and science classes. The online environment, however, affords us less time to state our exact positions

or develop as sophisticated of arguments before we are boiled down by others into deindividuated groups. Children must be prepared for this, as indeed so much of our face-to-face communication is, if not displaced by CMC, being supplemented by it. Let us not accept an inevitable cycle of toxic disinhibition, where rationality and civility become conflated for some, who then expect others—many with less privilege—to tolerate and work within uncivil discourse. Let us instead encourage introspection upon our own online presence, as this appears to be one way to combat toxic disinhibition without interrupting the boon to open deliberation afforded by benign disinhibition.

We are told that it is often wise to forego an argument and “agree to disagree.” But what happens when we cannot do this any longer? Does following this idiom actually prevent injury, or only put it off? What if those involved in the argument experience unequal consequences as a result of this non-action; for example, if “agreeing to disagree” had been the end of the discussion during any prominent Supreme Court case, such as those relating to women’s suffrage, abortion policy, abolishing slavery, allowing same-sex couples to marry, and so on—what then? Agreeing to disagree does not produce a moderate or central result—it only holds in place the conditions that benefit, often, those with the luxury to bow out of debate and remain comfortable.

In the task presented within this study, participants were prevented from agreeing to disagree, because doing so would not result a single set of policies, as they were directed to produce. When political debate is had for fun, we may be tempted to avoid getting to this point. However, when people of differing opinions are forced together on the same team and must yield a unified result, agreeing to disagree cannot the end point—the fight must be had, and ideally it must be had with rationality *and* civility. This

type of expectation will occur in many instances across one's life, whether they are working with a group for school or employment, acting as part of a team, making big decisions in intimate relationships, or crafting the very policies everyone will then be expected to live by. The online environment, then, with all of its potential for connecting us, more often and more rapidly, with teammates from all countries and time zones and interests, and where benign aspects of disinhibition may embolden us to ask for the outcomes we want, should be a space where we can practice agreeing to finish the fight—agreeing to agree.

REFERENCES

- Abramowitz, A. I. (2010). *The disappearing center: Engaged citizens, polarization, and American Democracy*. New Haven, CT: Yale University.
- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Al-Rawi, A. (2016). The online response to the Quran-burning incidents. In N. Mellor & K. Rinnawi (Eds.), *Political Islam and global media: The boundaries of religious identity* (105-121). New York, NY: Routledge.
- Amichai-Hamburger, Y., & McKenna, K. Y. A. (2006). The contact hypothesis reconsidered: Interacting via the Internet. *Journal of Computer-Mediated Communication*, *11*(3), 825-843. doi: 10.1111/j.1083-6101.2006.00037.x
- Bailenson, J. N., Yee, N., Blascovich, J., & Guadagno, R. E. (2008). Transformed social interaction in mediated interpersonal communication. In E. Konjin, S. Utz, M. Tanis, & S. B. Barnes (Eds.), *Mediated interpersonal communication* (77-98). New York, NY: Routledge.
- Barklage, R. (2016, August 26). Yes, I'm pro-life, but I'm not a stereotype. *Glamour*. Retrieved from <http://www.glamour.com/story/yes-im-pro-life-but-im-not-a-stereotype>
- Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture*, *5*(2), 171-180. doi:10.1037/ppm0000055
- Baym, N. K. (2010a). New forms of personal connection. In *Personal connections in the digital age* (pp. 1-21). Cambridge, UK: Polity Press.

- Baym, N. K. (2010b). New relationships, new selves?. In *Personal connections in the digital age* (pp. 99-121). Cambridge, UK: Polity Press.
- Behm-Morawitz, E. (2013). Mirrored selves: The influence of self-presence in a virtual world on health, appearance, and well-being. *Computers in Human Behavior*, 29(1), 119-128. doi:10.1016/j.chb.2012.07.023
- Bennett, J. (2014, August 29). Bubbles carry a lot of weight: Texting anxiety caused by little bubbles. *The New York Times*. Retrieved from <https://www.nytimes.com/2014/08/31/fashion/texting-anxiety-caused-by-little-bubbles.html>
- Bessièrè, K., Seay, A. F., & Kiesler, S. (2007). The ideal elf: Identity exploration in World of Warcraft. *CyberPsychology & Behavior*, 10(4), 530-535. doi: 10.1089/cpb.2007.9994
- Bohman, J. (2000). *Public deliberation: Pluralism, complexity, and democracy*. Cambridge, MA: MIT press.
- Bregman, A., & Haythornthwaite, C. (2003). Radicals of presentation: Visibility, relation, and co-presence in persistent conversation. *New Media & Society*, 5(1), 114-140.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, 55, 429-444.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, 28, 59-78.

Brown, A. (2017). What is so special about online (as compared to offline) hate speech?

Ethnicities. Retrieved from

https://ueaeprints.uea.ac.uk/64133/1/Accepted_manuscript.pdf

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun.

Personality and Individual Differences, 67, 97-102. doi:

10.1016/j.paid.2014.01.016

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality*

and Social Psychology, 63, 452-459. doi: 10.1037/0022-3514.63.3.452

Calhoun, C. (1988). Polulist politics, communication media and large scale societal

integration. *Sociological Theory*, 6(3), 219-241.

Callahan, S., & Callahan, D. (1987). Breaking through the stereotypes. In T. A.

Shannon, (Ed.), *Bioethics* (3rd ed.) (47-55). Mahwah, NJ: Paulist Press.

Cappella, J. N., & Jamieson, K. H. (1997). *Spiral of cynicism: The press and public good*.

New York, NY: Oxford University Press.

Caplan, B. (2007). *The myth of the rational voter*. Princeton, NJ: Princeton University

Press.

Carlton, C. L., Nelson, E. S., & Coleman, P. K. (2000). College students' attitudes toward

abortion and commitment to the issue. *The Social Science Journal*, 37(4), 619-

625.

Chen, G. M. & Ng, Y. M. M. (2016). Third-person perception of online comments: Civil

ones persuade you more than me. *Computers in Human Behavior*, 55, 736-742.

doi: 10.1016/j.chb.2015.10.014

- Chess, S., & Shaw, A. (2015). A conspiracy of fishes, or, how we learned to stop worrying about #Gamergate and embrace hegemonic masculinity. *Journal of Broadcasting & Electronic Media*, 59(1), 208-220. doi: 10.1080/08838151.2014.999917
- Citrin, J., Green, D. P., Muste, C., & Wong, C. (1997). Public opinion toward immigration reform: The role of economic motivations. *The Journal of Politics*, 59(3), 858-881.
- Cohen, D. L. (2014, January 17). Your practice – Financial advisers changing approaches for women. *Reuters*. Retrieved from <https://www.reuters.com/article/column-yourpractice/your-practice-financial-advisers-changing-approaches-for-women-idUSL2N0KP15720140117>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.
- Coles, A., & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior*, 60, 233-244.
- Cunningham, S. J., Nichols, D. M., Hinze, A., & Bowen, J. (2015, September). Digital news resources: An autoethnographic study of news encounters. In *International Conference on Theory and Practice of Digital Libraries* (pp. 84-96). Springer International Publishing.
- Dalton, R.J., & Wattenberg, M.P. (2000). *Parties without partisans: Political change in advanced industrial democracies*. Oxford, England: Oxford University Press.

- Dash, A. (2016, January 27). Against “Don’t read the comments”. Retrieved from <https://medium.com/humane-tech/against-don-t-read-the-comments-ae43ce515b9#.fj3caxcdw>
- D’Costa, K. (2013, July 29). Don’t read the comments! (Why do we read the online comments when we know they’ll be bad?). *Scientific American*. Retrieved from <https://blogs.scientificamerican.com/anthropology-in-practice/dont-read-the-comments-why-do-we-read-the-online-comments-when-we-know-theyll-be-bad/>
- Deegan, M. P., Hehman, E., Gaertner, S. L., & Dovidio, J. F. (2015). Positive expectations encourage generalization from a positive intergroup interaction to outgroup attitudes. *Personality and Social Psychology, 41*(1), 52-65. doi: 10.1177/0146167214556240
- Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, tasks, and communication processes: A theory of media synchronicity. *MIS Quarterly, 32*(3), 575-600.
- de Vreese, C. H., & Semetko, H. A. (2002). Cynical and engaged: Strategic campaign coverage, public opinion, and mobilization in a referendum. *Communication Research, 29*, 615-641.
- DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the Internet. *Annual Review of Sociology, 27*(1), 307-336. doi: 10.1146/annurev.soc.27.1.307
- Dir, A. L., Coskunpinar, A., Steiner, J. L., & Cyders, M. A. (2013). Understanding differences in sexting behaviors across gender, relationship status, and sexual

identity, and the role of expectancies in sexting. *Cyberpsychology, Behavior, and Social Networking*, 16(8), 568-574. doi: 10.1089/cyber.2012.0545

Douthat, R. (2013, January 26). Divided by abortion, united by feminism. *The New York Times*. Retrieved from <http://www.nytimes.com/2013/01/27/opinion/sunday/douthat-divided-by-abortion-united-by-feminism.html>

Drouin, M., Miller, D., Wehle, S. M. J., & Hernandez, E. (2016). Why do people lie online? "Because everyone lies on the Internet". *Computers in Human Behavior*, 64, 134-142. doi: 10.1016/j.chb.2016.06.052 0747-5632/©

Ellis, D. G., & Maoz, I. (2007). Online argument between Israeli Jews and Palestinians. *Human Communication Research*, 33(3), 291-309. doi:10.1111/j.1468-2958.2007.00300.x

Ertelt, S. (2011, June 6). Republicans strongly pro-life, Democrats strongly pro-abortion. *LifeNews.com*. Retrieved from <http://www.lifenews.com/2011/06/06/republicans-strongly-pro-life-democrats-strongly-pro-abortion/>

Eveland, W., Morey, A., & Hutchens, M. (2011). Beyond deliberation: New directions for the study of informal political conversation from a communication perspective. *Journal of Communication*, 61(6), 1082-1103. doi: 10.1111/j.1460-2466.2011.01598.x

Farrell, H. (2011). The consequences of the Internet for politics. *Annual Review of Political Science*, 15, 35-52. doi: 10.1146/annurev-polisci-030810-110815

Filipovic, J. (2014, November 20). Why more than 100 women are sharing their abortion stories. *Cosmopolitan*. Retrieved from

<http://www.cosmopolitan.com/politics/news/a33460/abortion-story-speakout/>

Fiorina, M. P., Abrams, S. J., & Pope, J. C. (2005). *Culture war? The myth of a polarized America*. New York, NY: Pearson Longman.

Flanagin, A. J., Tiyaamornwong, V., O'Connor, J., & Seibold, D. R. (2002). Computer-mediated group work: The interaction of sex and anonymity. *Communication Research, 29*(1), 66-93.

Fox, J., Cruz, C., & Lee, J. Y. (2015). Perpetuating online sexism offline: Anonymity, interactivity, and the effects of sexist hashtags on social media. *Computers in Human Behavior, 52*, 436-442. doi: 10.1016/j.chb.2015.06.024

Fox, J., & Moreland, J. J. (2015). The dark side of social networking sites: An exploration of the relational and psychological stressors associated with Facebook use and affordances. *Computers in Human Behavior, 45*, 168-176. doi: 10.1016/j.chb.2014.11.083

Gahagan, K., & Vaterlaus, J. M., & Frost, L. R. (2016). College student cyberbullying on social networking sites: Conceptualization, prevalence, and perceived bystander responsibility. *Computers in Human Behavior, 55*, 1097-1105. doi: 10.1016/j.chb.2015.11.019

Garrett, K. R. (2009). Echo chambers online? Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication, 14*(2), 265-285. doi:10.1111/j.1083-6101.2009.01440.x

- Garrett, R. K., Carnahan, D., & Lynch, E. K. (2013). A turn toward avoidance? Selective exposure to online political information, 2004-2008. *Political Behavior*, 35(1), 113-134. doi:10.1007/s11109-011-9185-6
- Gastil, J., Black, L., & Moscovitz, K. (2008). Ideology, attitude change, and deliberation in small face-to-face groups. *Political Communication*, 25(1), 23-46. doi:10.1080/10584600701807836
- Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Doubleday.
- Goffman, E. (1963). *Behavior in public places*. New York, NY: The Free Press.
- Govern, J. M., & Marsch, L. A. (2001). Development and validation of the situational self-awareness scale. *Consciousness and Cognition*, 10(3), 366-378.
- Gutmann, A., & Thompson, D. (1990). Moral conflict and political consensus. *Ethics*, 101(1), 64-88.
- Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, MA: MIT Press.
- Habermas, J. (1995). Reconciliation through the public use of reason: Remarks on John Rawls's political liberalism. *The Journal of Philosophy*, 92(3), 109-131.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834.
- Hall, J. A., Park, N., Song, H., & Cody, M. J. (2010). Strategic misrepresentation in online dating: The effects of gender, self-monitoring, and personality traits. *Journal of Social and Personal Relationships*, 27(1), 117-135. doi:10.1177/0265407509349633

- Halpern, D. & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 29(3), 1159-1168. doi:10.1016/j.chb.2012.10.008
- Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York, NY: The Guilford Press.
- Hermida, A. (2016). Social media and the news. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE handbook of digital journalism* (81-94). Thousand Oaks, CA: SAGE Publications Inc.
- Heron, M. J., Belford, P., & Goker, A. (2014). Sexism in the circuitry: Female participation in male-dominated popular computer culture. *Computers & Society*, 44(4), 18-29.
- Hess, J. A., & Rueb, J. D. (2005). Attitudes toward abortion, religion, and party affiliation among college students. *Current Psychology*, 24(1), 24-42.
- Hinkle, S., Taylor, L. A., Fox-Cardamone, D. L., & Cook, K. F. (1989). Intragroup identification and intergroup differentiation: A multicomponent approach. *British Journal of Social Psychology*, 28, 305-317.
- Hinnant, A., Subramanian, R., & Young, R. (2016). User comments on climate stories: Impacts of anecdotal vs. scientific evidence. *Climatic Change*, 138, 411-424. Doi: 10.1007/s10584-016-1759-1.
- Ho, S. S., & McLeod, D. M. (2008). Social-psychological influences on opinion expression in face-to-face and computer-mediated communication. *Communication Research*, 35(2), 190-207. doi:10.1177/0093650207313159

- Hughes, B. L., Ambady, N., & Zaki, J. (2017). Trusting outgroup, but not ingroup members, requires control: Neural and behavioral evidence. *Social Cognitive and Affective Neuroscience, 12*(3), 372-381. doi: 10.1093/scan/nsw139
- Hunt, D., Atkin, D., & Krishnan, A. (2012). The influence of computer-mediated communication apprehension on movies for Facebook use. *Journal of Broadcasting & Electronic Media, 56*(2), 187-202. doi: 10.1080/08838151.2012.678717
- Hutchens, M. J., Cicchirillo, V. J., & Hmielowski, J. D. (2014). How could you think that?!?!: Understanding intentions to engage in political flaming. *New Media & Society, 17*(8), 1201-1219. doi: 10.1177/1461444814522947
- Hwang, H., Kim, Y., & Huh, C. U. (2014). Seeing is believing: Effects of uncivil online debate on political polarization and expectations of deliberation. *Journal of Broadcasting & Electronic Media, 58*(4), 621-633. doi: 10.1080/08838151.2014.966365
- Inzlicht, M., & Kang, S. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology, 99*, 467-481.
- Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication, 59*(1), 19-39, doi: 10.1111/j.1460-2466.2008.01402.x
- Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science, 59*(3), 690-707. doi: 10.1111/ajps.12152

- James, C. (2014). *Disconnected: Youth, new media, and the ethics gap*. Cambridge, MA: The MIT Press.
- Jourard, S. M. (1971). *The transparent self*. New York, NY: Van Nostrand Reinhold.
- Kahan, D. M. (2007). The cognitively illiberal state. *Stanford Law Review*, 60(1), 115-154.
- Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, 2(10), 732-735.
doi:10.1038/nclimate1547
- Karpowitz, C. F., Mendelberg, T., & Shaker, L. (2012). Gender inequality in deliberative participation. *American Political Science Review*, 106(3), 533-547.
doi:10.1017/S0003055412000329
- Kelly, J., Fisher, D., & Sith, M. (2009). Friends, foes, and fringe: Norms and structure in political discussion networks. In T. Davies & S. P. Gangadharan (Eds.), *Online deliberation: Design, research, and practice* (83-93). Chicago, IL: CSLI Publications.
- Kercher, K. (1992). Assessing subjective well-being in the old-old: The PANAS as a measure of orthogonal dimensions of positive and negative affect. *Research on Aging*, 14(2), 131-168.
- Kersting, N., & Zimmermann, T. (2014). Online-comments: Deliberative or demonstrative political participation on the Internet? In *Conference for E-Democracy and Open Government* (p. 35).

- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, *39*, 1123-1134.
- Kim, N., & Wojcieszak, M. (2018). Intergroup contact through online comments: Effects of direct and extended contact on outgroup attitudes. *Computers in Human Behavior*, *81*, 63-72. doi: 10.1016/j.chb.2017.11.013
- Kroll, B. M. (2005). Arguing differently. *Pedagogy*, *5*(1), 37-60.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480-498.
- Kwon, K. H., & Gruzd, A. (2017). Is aggression contagious online? A case of swearing on Donald Trump's campaign videos on YouTube. Proceedings from *The 50th Hawaii International Conference on System Sciences*. Waikoloa, HI: HICSS.
- Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, *28*, 434-443. doi:10.1016/j.chb.2011.10.014
- Lapidot-Lefler, N., & Barak, A. (2015). The benign online disinhibition effect: Could situational factors induce self-disclosure and prosocial behaviors? *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *9*(2), article 3. doi: 10.5817/CP2015-2-3
- Lea, M., & Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing*, *2*, 321-341.
- Lee, H., & Pang, N. (2014). Responding to the haze: Information cues and incivility in the online small world. *Information Research*, *19*(4), 40-54.

- Levine, P. (2000). The Internet and civil society. *Philosophy and Public Policy*, 20(4), 1–8.
- Lieberman, M. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. *Psychological Bulletin*, 126(1), 109-137.
- Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge, MA: Cambridge University Press.
- Lord, C.S., Ross, L., & Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098–2109.
- Mackay, C. (1932). *Extraordinary popular delusions and the madness of crowds* (2nd ed.). Boston, MA: Page. (Original work published 1852)
- MacKuen, M., Wolak, J., Keele, L., & Marcus, G. E. (2010). Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science*, 54(2), 440-458. doi: 10.1111/j.1540-5907.2010.00440.x
- Massanari, A. (2015). #Gamergate and The Fappening: How Reddit’s algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 19(3), 329-346. doi: 10.1177/1461444815608807
- McKenna, K. Y. A., & Bargh, J. A. (2000). Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. *Personality and Social Psychology Review*, 4(1), 57–75.
- Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media. *Communication Research*, 41(8), 1042-1063.

- Michikyan, M., Subrahmanyam, K., & Dennis, J. (2014). Can I tell you who I am? Neuroticism, extraversion, and online self-presentation among young adults. *Computers in Human Behavior, 33*, 179-183. doi: 10.1016/j.chb.2014.01.010
- Misoch, S. (2015). Stranger on the Internet: Online self-disclosure and the role of visual anonymity. *Computers in Human Behavior, 48*, 535-541. doi: 10.1016/j.chb.2015.02.027
- Montoya-Weiss, M. M., Massey, A. P., & Song, M. (2001). Getting it together: Temporal coordination and conflict management in global virtual teams. *The Academy of Management Journal, 44*(6), 1251-1262.
- Moore, M. J., Nakano, T., Enomoto, A., & Suda, T. (2012). Anonymity and roles associated with aggressive posts in an online forum. *Computers in Human Behavior, 38*(3), 861-867. doi: 10.1016/j.chb.2011.12.005
- Moy, P., & Pfau, M. (2000). *With malice toward all? The media and public confidence in democratic institutions*. Westport, CT: Praeger.
- Mutz, D., & Martin, P. (2001). Facilitating communication across lines of political difference: The role of mass media. *American Political Science Review, 95*(1), 97-114. doi: 10.1080/0163853X.2011.577391
- Nguyen, M., Bin, Y. S., & Campbell, A. (2012). Comparing online and offline self-disclosure: A systematic review. *Cyberpsychology, Behavior, and Social Networking, 15*(2), 103-111. doi:10.1089/cyber.2011.0277
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220.

- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
- Nowak, K. (2001). Defining and differentiating copresence, social presence and presence as transportation. Proceedings from *Presence 2001 Conference*, Philadelphia, PA.
- Nyhan, B., & Reifler, J. (2015). Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine*, 33(9), 459-464. doi:10.1016/j.vaccine.2014.11.017
- Page, B. I. (1996). *Who deliberates? Mass media in modern democracy*. Chicago, IL: University of Chicago Press.
- Papacharissi, Z. (2002). The virtual sphere: The Internet as a public sphere. *New Media & Society*, 4(1), 9-27.
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259-283. doi: 10.1177/1461444804041444
- Parkin, S. (2014, October 17). Gamergate: Ascandal erupts in the video-game community. *The New Yorker*. Retrieved from <http://www.newyorker.com/tech/elements/gamergate-scandal-erupts-video-game-community>
- Parks, M. R., & Floyd, K. (1996). Making friends in cyberspace. *Journal of Computer-Mediated Communication*, 1(4). doi: 10.1111/j.1083-6101.1996.tb00176.x
- Parks, M., & Roberts, L. (1997). Making MOOsic: the Development of personal relationships on-line and a comparison to their off-line counterparts. Proceedings from *Western Speech Communication Association Conference*, Monterrey, CA.

- Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). "Us" and "them": Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology*, *59*, 475–486. doi: 10.1037/0022-3514.59.3.475
- Postmes, T., Spears, R., & Lea, M. (1998). Breaching or building social boundaries? SIDE-effects of computer-mediated-communication. *Communication Research*, *25*, 689-715.
- Price, V., Tewksbury, D., & Powers, E. (1997). Switching trains of thought: The impact of news frames on readers' cognitive responses. *Communication Research*, *24*, 481-506.
- Puri, K., & Mulay, P. (2016). Making machines talk with intelligent senses. *International Journal of Computer Science and Mobile Computing*, *5*(5), 296-306.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Rendon, M. (1984). Myths and stereotypes in minority groups. *International Journal of Social Psychiatry*, *30*(4), 297-309.
- Riles, J. M., Sangalang, A., Hurley, R. J., & Tewksbury, D. (2015). Framing cancer for online news: Implications for popular perceptions of cancer. *Journal of Communication*, *65*(1), 1018-1040. doi: 10.1111/jcom.12183
- Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., & Broverman, D. M. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology*, *32*, 287-295.

- Rourke, L., & Kanuka, H. (2007). Barriers to online critical discourse. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 105-106. doi: 10.1007/s11412-007-9007-3
- Ryan, T., & Xenos, S. (2011). Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. *Computers in Human Behavior*, 27(5), 1658-1664. doi:10.1016/j.chb.2011.02.004
- Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, 8(1), 18-33. doi:10.1080/17512786.2013.813194
- Shudson, M. (1995). *The Power of News*. Cambridge, MA: Harvard University Press.
- Sears, D. O. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, 31(2) 194-213. doi:10.1086/267313
- Settle, J. E., Bond, R. M., Coviello, L., Fariss, C. J., Fowler, J. H., & Jones, J. J. (2016). From posting to voting: The effects of political competition on online political engagement. *Political Science Research and Methods*, 4(2), 361-378. doi:10.1017/psrm.2015.1
- Shim, J. W., & Paul, B. M. (2014). The role of anonymity in the effects of inadvertent exposure to online pornography among young adult males. *Social Behavior and Personality*, 42(5), 823-834. doi: 10.2224/sbp.2014.42.5.823
- Smith, M. C. (2016, April 19). I'm a Republican and I am pro-choice. *Odyssey*. Retrieved from <https://www.theodysseyonline.com/republican-pro-choice>
- Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology*, 15(4), 330-342.

- Son, L. K., Kornell, N., Finn, B., & Cantlon, J. F. (2012). Metacognition and the social animal. In P. Briñol, & K. DeMarree. (Eds.), *Social metacognition* (159-175). New York, NY: Taylor & Francis Group.
- Spears, R., & Lea, M. (1992). Social influence and the influence of the “social” in computer-mediated communication. In M. Lea (Ed.), *Contexts of computer-mediated communication* (30-65). London, England: Harvester-Wheatsheaf.
- Stein, J. (2016). Tyranny of the mob. *Time*, 188(8), 26-32. Retrieved from <http://coramcivic.org/resources/TYRANNY+OF+THE+MOB.pdf>
- Stephan, W. G., & Stephan, C. W. (1984). The role of ignorance in intergroup relations. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (229-255). New York, NY: Academic Press.
- Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, 7(3), 321-326. doi:10.1089/1094931041291295.
- Suler, J. (2008). Image, word, action: Interpersonal dynamics in a photo-sharing community. *CyberPsychology & Behavior*, 11(5), 555-560. doi 10.1089/cpb.2007.0153
- Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information*, 13, 65– 93.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

- Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology, 38*(2), 227-242.
- Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know each other one bit at a time. *Human Communication Research, 28*(3), 317-348.
doi:10.1111/j.1468-2958.2002.tb00811.x
- Trepte, S., & Loy, L. S. (2017). Social identity theory and self-categorization theory. *The international encyclopedia of media effects*. doi:
10.1002/9781118783764.wbieme0088
- Tuchman, B. W. (1984). *The march of folly: From Troy to Vietnam*. New York, NY: Ballantine Books.
- Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. In N. Ellemers, R. Spears, & B. Dossje (Eds.), *Social identity: Context, commitment, content* (6–34). Oxford, UK: Blackwell.
- Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. *Computers in Human Behavior, 41*, 253-261. doi:10.1016/j.chb.2014.09.036
- Valentino, N. A., Beckmann, M. N., & Buhr, T. A. (2001). A spiral of cynicism for some: The contingent effects of campaign news frames on participation and confidence in government. *Political Communication, 18*, 347-367.

- Valkenburg, P. M., & Peter, J. (2011). Online communication among adolescents: An integrated model of its attraction, opportunities, and risks. *Journal of Adolescent Health, 48*, 121-127. doi:10.1016/j.jadohealth.2010.08.020
- “Views about abortion by state.” (2014). *Pew Research Center*, Washington, D.C.
Retrieved from <http://www.pewforum.org/religious-landscape-study/compare/views-about-abortion/by/state/>
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23*(3), 3-43.
doi:10.1177/009365096023001001
- Walther, J. B., Hofer, E., Ganayem, A., & Shonfeld, M. (2015). Computer-mediated communication and the reduction of prejudice: A controlled longitudinal field experiment among Jews and Arabs in Israel. *Computers in Human Behavior, 52*, 550-558. doi: 10.1016/j.chb.2014.08.004
- Walton, S. C., & Rice, R. E. (2013). Mediated disclosure on Twitter: The roles of gender and identity in boundary impermeability, valence, disclosure, and stage. *Computers in Human Behavior, 29*, 1465-1474. doi:10.1016/j.chb.2013.01.033
- Warkentin, D., Woodworth, M., Hancock, J. T., & Cormier, N. (2010). Warrants and deception in computer-mediated communication. In K. Inkpen (Ed.), *Proceedings of the 2010 ACM conference on computer supported cooperative work* (pp. 9e12). New York, NY: ACM.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

- Wilhelm, A. G. (1998). Virtual sounding boards: How deliberative is on-line political discussion?. *Information Communication & Society, 1*(3), 313-338.
- Williams, E. (1977). Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin, 84*(5), 963-976.
- Winter S., Brückner C., & Krämer, N.C. (2015). They came, they liked, they commented: Social influence on Facebook news channels. *Cyberpsychology, Behavior, and Social Networking, 18*, 431–436.
- Wojcieszak, M., & Mutz, D. (2009). Online groups and political discourse: Do online discussion spaces facilitate exposure to political disagreement? *Journal of Communication, 59*(1), 40-56. doi:10.1111/j.1460-2466.2008.01403.x

VITA

Michelle E. Funk attended Indiana University's Media School for her undergraduate (2012) and master's (2015) degrees before joining the University of Missouri's Department of Communication for her doctoral degree (2018). Her primary research interest lies in media and identity, particularly in online contexts. Early in her graduate research, she focused on sexually explicit media and sexual expectations between genders, including the influence of pornography consumption on attitudes toward affirmative action for women. The results of this focus have been published in *Psychology of Women Quarterly* and *Archives of Sexual Behavior*. Her specific interest is in the examination of mediated conflicts between ingroups and outgroups comprised of different social identities and hegemonic positions. Beyond studying gender and media sexuality, she has since conducted research that examines intersectional identities of ethnicity, class, and religiosity.

Most recently, she has been interested in investigating the impact of political social groups as they impact social cognitive processes used when making candidate evaluations based off of media stimuli. One such study has been published in *Communication Studies*, with others at various stages of preparation and submission. She is excited to use the present dissertation as a launching-off point for a trajectory of scholarship related to online deliberation, outgroup relations, and educational initiatives related to communication in younger populations.

Though the majority of her research is conducted from an empirical, post-positivist lens, she has also written about the need for a stronger dialogue between quantitative and critical methodologies, as researchers from all epistemological

standpoints rely on one another to reveal different facets of cultural phenomena. She is deeply interested in critically thinking about the ethics and practices of data collection, including data deriving from social media within the uniquely complex online context.