UNIVERSITY OF CALIFORNIA

Santa Barbara

An Empirical Test of the Effects of Political Correctness: Implications for Censorship, Self-Censorship, and Public Deliberation

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Communication

by

Becky R. Ford

Committee in charge:

Professor Scott Reid, Chair

Professor Howie Giles

Professor Jennifer Kam

December, 2017

The dissertation of Becky R. Ford is approved.

_____

Howard Giles

_____

Jennifer A. Kam

_____

Scott A. Reid, Committee Chair

November 2017

An Empirical Test of the Effects of Political Correctness: Implications for Censorship, Self-Censorship, and Public Deliberation

ACKNOWLEDGEMENTS

This dissertation and my five years of grad school would not have been possible without the help of many people. To my advisor, Scott Reid, thank you for always allowing me to pursue my own interests, even when they do not align with yours, and for giving me a strong foundation on which to build my career. To my committee members Howie Giles and Jennifer Kam, thank you for the all of the stimulating debate during defenses, for challenging me to think about my research questions from multiple angles, and for your general support through this dissertation process.

To all of my friends at colleagues at UCSB, it was much easier to get through grad school knowing you had my back. To my research assistants, Natasha Mehrdad, Daniella Mikhael-Fard, Bowen Xie, and Kat Keller, I could not have accomplished this project without your diligent data collection. To Dana Mastro, thank you for your feedback on this dissertation and on other projects throughout grad school. To Amy Robinson, Renee Robinson, Brooke Hundtoft, Laura Jones, and Jamie Worsek, thank you for not only putting your red teacher-pens to use by proofreading drafts of this dissertation, but more importantly, for pretending to be happy about it. To my parents, thank you for always encouraging us to pursue education, and for never pressuring me to hurry up and get a "real" job. And last, to Sean, thank you for being my support system, be it emotional support, tech support, or whatever other support I needed, and for never complaining when you "had" to spend the weekend playing video games because I needed to work. I never could have made it through grad school without you.

VITA OF BECKY R. FORD
November 2017

EDUCATION

Bachelor of Arts, University of Illinois, Urbana-Champaign, June 2008

Master of Arts in Communication, University of California, Santa Barbara, June 2014

Doctor of Philosophy in Communication, University of California, Santa Barbara, September 2017

PROFESSIONAL EMPLOYMENT

2016-2017	Graduate Student Researcher Department of Communication, University of California Santa Barbara
2016	Teaching Associate Instructor of record Department of Communication, University of California Santa Barbara
2012-2016	Teaching Assistant Department of Communication, University of California Santa Barbara

PUBLICATIONS

Ford, B.R. & Stohl, C. (forthcoming, Spring 2018). Corporate social responsibility and product reviews: A longitudinal content analysis. *Journal of Brand Management.*

Ford, B.R. (in press). Political correctness. An invited chapter for the *Oxford Encyclopedia of Intergroup Communication*.

Robinson, B. & Coveleski, S. (online first). Don't say that to ME: Opposition to targeting in weight-related intervention messages. *Health Communication*. doi: 10.1080/10410236.2016.1250189

Robinson, B. & Reid, S.A. (2016). Political correctness. In H. Giles & A. Maass (Eds.) *Advances in Intergroup Communication*. (pp. 85-100). New York, NY: Peter Lang Publishing.

Reid, S.A. & Robinson, B. (2015). Social identity theories. In Berger, C. & Roloff, M. *International Encyclopedia of Interpersonal Communication*. (pp. 1-6). Wiley-Blackwell.

ABSTRACT

An Empirical Test of the Effects of Political Correctness: Implications for Censorship, Self-Censorship, and Public Deliberation

by

Becky R. Ford

For over 30 years, scholars, journalists, and politicians have debated the costs and benefits of Political Correctness (PC). Those who support PC claim that it benefits historically disadvantaged groups by protecting them from discrimination and encourages diverse representation. Opponents to PC claim that it inhibits freedom of expression and thus public deliberation. However, despite three decade of debate, PC is undertheorized and has received little empirical investigation.

In this dissertation, following theorizing by Robinson and Reid (2016a), and research on social identity, self-categorization, and public deliberation, I propose that PC is rooted in identity politics (on the right and left), and should be viewed as a tool to control discourse in intergroup conflict. This dissertation argues that PC is an ideology that (among other things) relies on adherents' perceptions of protected and perpetrator groups, involves the imposition of social sanctions and censorship, and justifies such actions by appealing to the moral failings of actors whose actions are judged anti-PC (e.g., sexist or racist). Further, individuals who believe that their views and actions may be perceived as anti-PC may be more likely to

self-censor to avoid sanctions and being judged as immoral. This may ultimately affect public deliberation due to a decreased tolerance of diverse viewpoints.

The relationships between PC, perceptions of victimhood, support for censorship, self-censorship, and public deliberation were tested across three studies. The survey findings from Study 1 indicate that participants' perceptions of victimhood were predicted by their political ideology, such that the more liberal participants were the more likely they were to perceive victimization among racial, sexual, and religious minorities, and the more conservative participants were, the more likely they were to perceive victimization among Whites, Christians, and males. The same effects of political ideology were observed for support for censorship of political opponents. Study 2 primed participants using either PC code words or a control that did not include code words, and found that the more conservative participants were, the more likely they were to report self-censoring, but only after exposure to a PC prime. Study 3 had participants engage in an online conversation with a confederate under either a PC or non-PC prime. Participants exposed to the PC prime argued with lower levels of integrative complexity (a measure of the extent to which people recognize alternative view points) than those in a non-PC condition, and the more liberal participants were, the less of integrative complexity they exhibited. Taken together, these studies confirm that PC involves competition between liberals and conservatives, that PC norms produce self-censorship among moderates and relatively conservative students, and produce less cognitively complex reasoning about political subjects, particularly among liberals.

TABLE OF CONTENTS

Chapter I: Political Correctness

Political correctness (PC) has grown in cultural and political significance in Western countries since the 1980s (Berman, 1992; Chait, 2015; Ford, 2017; Hughes, 2010; Robinson & Reid, 2016a; Weigel, 2016; Wilson, 1995). Without possessing a shared definition of PC, political commentators and academics (primarily in the humanities) have argued over its social costs and benefits. The most commonly referenced argument by those who endorse PC is that it prevents discrimination against historically disadvantaged groups, and that this group-protective function is a social good that promotes fairness and equality (Fairclough, 2003; Fish, 1992; Whitney & Wartella, 1992). On the other hand, scholars, journalists, and politicians (on both the political left and right) have argued that PC has the potential to decrease free expression and thereby undermine science and democracy (Berman, 1992; Al-Gharbi & Haidt, 2015; Chait, 2015; D'Souza, 1992; Loury, 1994). While each of these perspectives likely contains some truth, many questions regarding the phenomenology, precipitating conditions, and social consequences of PC remain unexamined by behavioral science.

PC is a multi-faceted phenomenon that involves political debate over discrimination and how best to deal with it (e.g., racism, sexism, hetero-sexism), highly valued social and political identities, normative prescriptions regarding "appropriate" and "inappropriate" language use, a sense of moral legitimacy about the way society ought to be, and endorsement of political policies and actions involving diversity and representation of social groups (e.g., affirmative action). It is likely that the phenomenological complexity of PC has hindered the development of a shared definition. Nonetheless, a case will be made that PC is rooted in intergroup conflict between people with opposed political identities (i.e., that it is a

form of identity politics), and that it is ideological in nature (i.e., it involves the promotion of broad-based norms across political affiliations). In elaborating this argument, Robinson and Reid (2016a) proposed that PC does not refer to any specific ideology or norm content, nor is PC the sole domain of the political left or right. Rather, PC enables those on the political right and left to engage in intergroup competition by controlling discourse, and by extension, political opponents. As will be demonstrated, this common thread is found throughout political conflict, and intergroup conflict more generally.

A key claim made regarding contemporary PC is that it has a chilling effect that undermines science and rational deliberation, particularly in (but not limited to) academia (e.g., Dreger, 2014; Kipnis, 2015; Lukianoff & Haidt, 2015; Schlosser, 2015). But what does this "PC university" look like, and how can it have these alleged effects on college campuses? Examples elicit the label "PC" include university administrators' decisions to prevent guest speakers from coming to campus if they are thought to espouse viewpoints that may offend some groups, or classroom policies wherein professors encourage the use of PC language. Those who oppose PC often cite examples similar to those just given to support their claim that PC has negative effects on university life. However, arguments in favor of and opposed to PC are yet to be subjected to rigorous empirical testing. The goal of this dissertation is to empirically test the conditions under which PC may lead to self-censorship and how this is related to rational deliberation.

This dissertation focuses on the intergroup relations that underpin and regulate PC, and the conditions under which PC norms might affect self-censorship. In this chapter I explain the historical context in which PC emerged, describe the phenomena in detail, and explicate the intergroup functions that PC serves. Chapter II focuses on the relationship

between PC and self-censorship, which is arguably a key concern because it is free expression that enables both majority and minority groups to promote their interests via the democratic processes (Al-Gharbi & Haidt, 2017). Last, I outline how intergroup relations can provide a better explanation for the connection between PC norms and self-censorship than current approaches (e.g., Spiral of Silence; Noelle-Neumann, 1974). The remaining chapters report three studies that test the connection between PC, perceptions of group victimhood, and endorsement of censorship (Study 1), PC and self-censorship (Studies 2 & 3) and rational deliberation (Study 3). This dissertation concludes with a discussion of its theoretical and practical implications and future areas of study on PC.

History of the PC Movement

The use of the term PC dates back to the 1930s in Moaist China and Stalinist Russia. The term meant to literally have correct politics, or to toe the party line. The phrase continued to be used through the 1960s by liberals, often ironically, to refer to members of their own party who they felt were acting too radically or self-righteously (Hughes, 2010; Weigel, 2016). Eventually this ironic use of the term by liberals to refer to ingroup struggles was co-opted by members of the right, and by the 1980s, "PC", as an abbreviated form, began to be used as a pejorative by conservatives to disparage the concerns of those promoting reforms of curricula and speech codes on college campuses.

PC has been, and continues to be, an issue that is primarily discussed in reference to universities (Berman, 1992; Hughes, 2010; Lukianoff & Haidt, 2015; Weigel, 2016). College campuses in the 1990s were rife with debate about what higher education should look like, and these debates often included discussion of PC. Several popular books (including D'Souza's *Illiberal Education*, Bloom's *The Closing of the American Mind,* and

Kimball's *Tenured Radicals*) claimed that those educated in the 1960s had become left-wing faculty (often referred to as "radicals"), and these "radicals" had created a dangerous PC climate on college campuses. These books shared the argument that many these "radical" faculty members possessed a PC agenda that promoted diversity and the protection of marginalized groups at the cost of academic standards and freedom. Those who supported PC ideology (whether or not they used that label) argued that diversifying universities was beneficial to all members of campus communities, as it fostered a diversity of perspectives that all students could learn from, and that it was essential to protect minorities from hate speech and other communication to which they might take offense if they were to feel like they belonged in academia (e.g., Berubé, 1992; Gates, 1992; Whitney & Wartella, 1994)[1].Though debates over PC has ebbed and flowed over the last 30 years (Weigel, 2016), many of these arguments remain unresolved, and continue inside and outside of academia. The current debate has tended to move away from diversity in representation to a focus on protecting minority groups on campus (e.g., debates over free expression, safe spaces, microaggressions, and trigger warnings) (Chait, 2015; Lukianoff & Haidt, 2015).

Political debates regarding PC have involved political positions on a range of issues, but a common thread has been debate within and between groups that have diverging political interests. The term itself can be problematic, because those on the left who are labeled PC (e.g., those who are pursuing social justice and diversity) typically do not identify as "PC," (Weigel, 2016). Similarly, those on the right frequently impose the term on liberals, and do

[1]The purpose of this dissertation is not to argue against the ideals of reducing inequality and increasing diversity. These are important goals. Instead, the purpose of this dissertation is to examine how PC is embedded in identity politics (as are many ideologies that have less socially desirable ideals), to reflect on how these ideologies follow many of the same

not apply it to their own political interests (e.g., conservatives who are offended by people who say "Happy Holidays" instead of "Merry Christmas" are rarely labeled "PC" nor would they consider the term self-descriptive). Nonetheless, the political right also engages in ideological struggles with their political opponents, and those actions are phenomenologically equivalent to the actions of those on the political left that frequently elicit the term "PC". For these reasons, it is important to be explicit about the phenomenology of PC, to consider how it describes the tactics of those on the left and right, and how it is currently playing a role in contemporary cultural and political life.

Features of Contemporary PC

PC is not a unitary phenomenon. After reviewing the literature, including journal articles, scholarly books, and news articles, and in line with Robinson and Reid (2016a), I identified several commonly referenced features of PC. In what follows, I describe these features and explain their relevance to defining PC as well as their importance in the current PC debate (see Table 1 for a list of features with examples).

Table 1

Features of PC phenomenology

Feature	Description	Examples
Protected & perpetrator groups	Prescriptions of ideology target specific protected groups (coalitional partners) and point to other groups as perpetrators (political opponents)	Affirmative action; Speech codes that prohibit speech against certain groups but not others

patterns that are seen in other areas of intergroup relations, and test how PC is related to self-censorship and public deliberation.

Language	Normative language use indicate support for an ideology and change attitudes towards groups or policies	Liberal: *non-binary*, *developmentally challenged* Conservative: *welfare queen, family values*
Conflicting norms	Interrelated prescriptions that may conflict with each other	Liberal: Conflicting views over which groups it is acceptable to criticize Conservative: Advocating for religious freedoms for Christians but not Muslims
Attributions of offense	Differing interpretations over what is and is not a violation of PC, thereby allowing people to strategically take offense	Robinson and Reid (2015) that group membership of the source of a comment influences how offensive participants find the comment to be
Social influence	People who do not support PC are often coerced into abiding by its prescriptions	A student at University of Michigan expressed that his religion viewed homosexuality as immoral, and was forced to write an apology letter in the school newspaper
Censorship	PC-dissenters are often sanctioned by being silenced	Milo Yiannopulos and Anne Coulter were prevented from speaking at UC Berkeley because of violent protests Keeanga-Yamahtta Taylor canceled her speaking tour due to death threats she received after criticizing the president

Protected and Perpetrator Groups

PC ideologies include norms that target certain groups for protection and identify others as threatening. For example, liberal PC has long rested on the argument that some groups are unfairly disadvantaged, that these groups are deserving of efforts to increase their

status and reduce the prejudice against them (Berman, 2992; Hughes, 2010; Lukianoff & Haidt, 2015). Current PC concerns make distinctions between protected groups that are deserving of increased representation (Blacks, Hispanics, women, sexual minorities, Muslims, and Native Americans; see Haidt & Leo, 2016), and perpetrator groups (typically White, heterosexual, Christian males, but also members of protected categories who do not toe the party line, such as People of Color who are accused of "acting White") that are typically considered responsible for victimizing those protected groups. It is true that these protected groups have been subjected to sustained discrimination. However, other groups, including Asians, Jews, Christians, the homeless, and poor disenfranchised Whites, often do not receive protected status, despite evidence of being discriminated against at times (Haidt & Leo, 2016).

In the 1990s, many of the arguments surrounding PC and protected classes involved issues of diversity of representation in hiring and curricula. This was often referred to as *multiculturalism*. Though there is disagreement over exactly what multiculturalism is, it is most often described as "treating society as the home of several equally valuable but distinct racial and ethnic groups" (Stimpson, 1992, p. 43). Starting in the 1980s, there was a push by some literature professors to change campus curricula to require students to take courses that expose them to non-Eurocentric literature (Berman, 1992). These professors argued that universities are supposed to teach students to appreciate the world's great texts, but universities had long overlooked literature that was not written by White men, meaning there were many great texts left out of college curricula. For example, why was it that William Shakespeare and John Steinbeck were the authors of great texts that *must* be read by college students, but those by Chinua Achebe and Toni Morrison were not essential reading? In

addition, those in favor of diversifying the authorship of assigned texts argued that because the US was becoming a more diverse country, it was important to empower members of historically powerless groups by highlighting the texts written by members of their groups (Gates, 1992; Stimpson, 1992). Berubé (1992), for example, argued that students can learn a great deal about other cultures from reading texts written by authors from those cultures, even if those texts are from outside of the traditional canon. In addition to the attention paid to course curricula, those who subscribed to the multiculturalist perspective on college campuses argued that women and People of Color were left out of the educational setting, and that this could be remedied, at least in part, by considering diversity as one factor in hiring decisions and student admissions (Stimpson, 1992).

Though issues of representation remain to be important debates on college campuses (e.g., an affirmative action case from University of Texas made it to the Supreme Court in 2016; under the Trump administration, the Justice Department said it would investigate the legality of Affirmative Action), arguments about representation are no longer the primary focus of the PC debate. Instead, much of the debate focuses on preventing protected groups from threats and being exposed to speech that they might consider offensive (Chait, 2015; Lukianoff & Haidt, 2015). Though these arguments used to be primarily about race and gender (Hughes, 2010), more recently arguments have moved from issues of race and ethnicity to sexuality, religion, and immigration status. Despite this shift in focus, I argue that liberal PC continues to promote the protection of certain protected groups of people more than others, and continues to point to groups that are likely to be political rivals (White men) as perpetrators. An in-depth discussion of this topic appears in Chapters II and III.

Language

PC ideologies have prescriptions that promote normative language that reflects their political preferences (e.g., preferred labels for protected groups) and aids their political goals (Hughes, 2010; Robinson & Reid, 2016a; Weigel, 2016). Language use has power because the words that people use do not merely reflect status and power relations, but actually aid in the creation of power where there was none, and obfuscates acts of power that are more effective when concealed (see Ng & Bradac, 1993; Reid & Ng, 1999, Sutton, 2010). Researchers have raised the possibility that changing language can create social change (e.g., Fairclough, 2003; Maass, Suitner, & Merkel, 2013). According to Fairclough (2003), when people use PC language (to use his example, "woman" instead of girl to refer to adult females), they are likely to internalize that change, which can eventually help raise the status of women (see Spender, 1980).

There is empirical support for this hypothesis. For example, immigrant groups in the United States that attract less complex and more negative ethnophaulisms (e.g., about physical traits, personality, work ethic, etc.) are less likely to marry U.S. citizens, are more likely to live in segregated neighborhoods, face employment exclusion, and commit suicide (Mullen & Rice, 2003; Mullen & Smyth, 2004). In addition, when people are exposed to gender neutral language versus gendered language (e.g., chairperson rather than chairman/chairwoman), they are less likely to endorse negative stereotypes about women (see Gabriel & Gygax, 2016 for a review). There is also evidence that replacing derogatory labels with a non-derogatory label (e.g., "fag" vs. "gay") has similar positive consequences (Carnaghi & Maass, 2007). Furthermore, relatively subtle discriminatory communication can harm those who are the targets of such language. For example, subtle language can produce stereotype threat, a phenomenon wherein being reminded of the stereotype about one's group

can cause the person to enact those stereotypes (e.g., Schmader, Johns, & Forbes, 2008). All of these findings suggest that language both reflects and constitutes, at least in part, the relationships between social groups (Sutton, 2010), supporting the PC claim that changing language can be beneficial in promoting group equality.

Research also shows that conservative language affects perceptions and political attitudes. For example, research has examined conservative politicians' use of racially coded language, used to prime racial fears among conservative voters without explicitly invoking race (Hurwtiz & Peffley, 2002, 2005; Mendelberg, 2002; Nelson & Kinder, 1996). In one study, Hurwitz and Peffley (2005) found that people who already harbored negative stereotypes about Black people were more likely to support spending money on prisons (vs. anti-poverty programs) as a solution to crime when they were primed with the code word "inner city" than when they were not. Essentially, these code words implicitly prime race for those who already hold negative racial stereotypes and cause greater support for harsher policies. These observations show that both the left and right uses language to promote endorsement of favored ideologies regarding race relations, (an in depth discussion of PC and language priming is found in Chapter IV).

Conflicting Norms

PC ideologies prescribe normative language, attitudes, and behaviors. However, these prescriptions may conflict with each other. For example, liberal PC proscribes stereotyping and discrimination, and also embraces many non-Western cultures. It may be considered politically incorrect (i.e., discriminatory) to derogate Islam and the cultures of some predominantly Muslim countries. However, at the same time, many of these predominantly Muslim countries have government policies and social norms that discriminate against

women and members of sexual minorities, and yet commentators who argue against these injustices may be excluded from political discussion for being discriminatory against Islam and Muslims. For example, in 2014, Ayaan Hirisi Ali, a woman and former Muslim who is an outspoken critic of sexism in the Muslim world, was supposed to receive an honorary degree at Brandeis University for her dedication to women's rights. However, her honorary degree and speaking engagement were cancelled when a group of students protested Ali's anti-Muslim stance (Pérez-Peña & Vega, 2014). Similar issues occur on the right. For example, conservatives often cite the First Amendment's protection of religious liberties when arguing for religious instruction in schools (e.g., both biology and intelligent design), but at the same time oppose public expression of other religious groups. Though there is little empirical work on how people manage conflicting norms and ideological prescriptions, it is clear that it makes it difficult for people to know exactly how to navigate and avoid violating PC, and allows for the strategic interpretation of norm violations.

Attributions of Offense

Another feature of PC is that when a PC norm is violated, people are often offended (or claim to be), ostensibly because a violation of PC is interpreted as threatening and/or discriminatory (e.g., racist, sexist, homophobic and so on) (Lukianoff & Haidt, 2015). However, just as some people may be offended by the use of profanity and others may not be, and some may be offended by communication about deities (e.g., depictions of the Prophet Mohammed, "taking the Lord's name in vain"), different people have different interpretations about when communication about marginalized groups is offensive. So how do people determine what communication is offensive? According to research on motivated reasoning, people desire to perceive situations objectively and accurately, but they are more

likely to reach a conclusion that is consistent with their preferences (see review in Ditto, Pizarro, & Tennenbaum, 2009). If message receivers are engaging in motivated reasoning (and research on the subject suggests that people do this most of the time), then offense taking is unlikely to be a purely logical response to the content of a communicative message, but is coupled with the receiver's prior attitudes, social identities, and ideological commitments, along with their reaction to the source of that message (Robinson & Reid, 2015; 2016a). Work on the Intergroup Sensitivity Effect (ISE) (e.g., Esposo, Hornsey, & Spoor, 2013; Hornsey & Imani, 2004; Hornsey, Trembath, & Gunthorpe, 2004), shows that people are more likely to take offense to criticism from outgroup members than ingroup members. Other research shows that offense taking is often on behalf of other groups, who may not actually be offended themselves. For example, Leets and Giles (1997) found that White participants interpreted explicit derogatory comments about Asian Americans as more harmful than did Asian Americans, but Asian American participants interpreted indirect derogatory comments to be more harmful than did Whites.

In addition to affecting how people interpret explicitly derogatory comments, these strategic attributions of offense have been found in how people interpret comments that are not explicitly derogatory, but nonetheless fall into a gray area of communication that *might* be offensive. Take, for example, three high profile cases of people who have been the recipients of harsh social sanctions for making arguments about biological sex differences in intelligence. In 2005, Larry Summers, then president of Harvard, was forced to step down from his position for citing research that attributed some, but not all, of gender inequality in math and science to biological explanations. Similarly, also in 2005, Helmuth Nyborg, a tenured professor, published research about sex differences in intelligence, and his opponents

accused him of scientific misconduct. Nyborg was almost fired from his job, except that an independent scientific committee could not find any significant methodological flaws in his research. More recently, outside of the university setting, James Damore was fired from his position as software engineer at Google for sending out a memo in which one of the many arguments he made advocated that Google's efforts to recruit more female engineers were misguided and sexist. In the cases of Summers, Nyborg, and Damore, not everyone interpreted their comments as offensive and worthy of sanction, and yet in all three cases, those who were offended were able to successfully call for sanctions. These examples highlight the fact that offense is an attribution rather than an objective evaluation, and that those who take offense may be engaging in motivated reasoning in making their interpretations.

In testing the extent to which offense and the decision to sanction offenders may be strategic, Robinson and Reid (2015) found that people are likely to interpret information along intergroup lines. In their study, the researchers presented participants with a fabricated newspaper article (modeled on Nyborg's 2005 study) in which a researcher provided evidence that men have bigger brains than women, and claimed that men therefore have higher general intelligence than women. Robinson and Reid found that female participants were more likely to find the researcher offensive than male participants. Participants were also more offended when the researcher was presented as a man than when the researcher was presented as a woman. Interestingly, the more a participant desired a high status position, the more likely they were to endorse sanctions, including censoring the speaker from future presentations, when they thought the speaker was high (vs. low) status (e.g., a professor versus graduate student). These findings suggest that offense taking is often

subjective, and can reflect reactions to the source of a comment and political considerations rooted in social identities as much as the content of the message (Loury, 1994; Robinson & Reid, 2016a).

Social Influence

An important feature of PC is that those who endorse the ideology try to influence others to abide by its prescriptions and proscriptions. There are many people who endorse PC because they are persuaded by its truth and authority (Robinson & Reid, 2016a). When people hold attitudes consistent with PC's prescriptions, they are likely to follow PC norms (e.g., using PC language) because they are persuaded that it is the right thing to do. However, not all people who abide by PC's prescriptions actually hold PC consistent attitudes. Because infractions against PC norms can lead to quite severe social sanctions (e.g., public shaming, job termination), many people may comply with PC norms only to avoid such sanctions, (Chait, 2015; D'Souza, 1992; Ford, 2017; Kipnis, 2015; Robinson & Reid, 2016a). For example D'Souza (1992) describes a case at the University of Michigan in which a student expressed his religious belief that same-sex relationships were immoral, and was then was forced by the university administration to write an apology letter in the school newspaper (D'Souza & MacNeil, 1992). D'Souza argued that "a subject that is a matter of legitimate discussion is, instead, settled not by reference to persuasion but by reference to force" (p. 35). This remains a common argument against PC. Similar social sanctions occur when those on the left violate conservative ideologies. For example, Lisa Durden, an adjunct professor at Essex College, was fired from her job after receiving criticism for her comments on Fox News, where she argued in defense of restricting participation in a Black Lives Matter rally to Black people. When people adhere to PC norms (whether they are liberal or

conservative) because of fear of sanctions (e.g., expressing normative attitudes to staying silent, using normative language), it often does not change their underlying attitudes, and can result in the false impression that fewer people hold prejudices than is actually the case (Avery & Steingard, 2008).

It is certainly possible that sanctioning PC violations can have positive consequences. For example, proscribing hate speech may have positive effects for the overall campus climate, regardless of whether or not those who would have otherwise expressed hate speech actually change their attitudes. However, another possibility is that using sanctions to enforce norms may produce a boomerang effect. For example, there is evidence that people who have a high (vs. low) external motivation to avoid expressing prejudice (i.e., a desire to avoid sanctions when being perceived as prejudiced) are more likely to retain negative racial stereotypes in interracial interactions, and have higher levels of implicit racial bias (Plant & Devine, 2001). Those who are externally motivated to avoid expressing prejudice, rather than internally motivated (i.e., persuaded that prejudice is wrong), are also likely to behave towards outgroup members in line with their true beliefs, rather than consistent with their publicly expressed views (Plant, Devine, & Brazy, 2003). This is perhaps why even those generally in support of liberal PC ideology have raised fears that enforcing PC on those who do not endorse it can make it appear that discrimination is less of a problem than it really is (e.g., Avery & Steingard, 2008; Ehrenreich, 1992; Sutton, 2010).

Self- and Other-Censorship

Though there are many sanctions that may be levied against those who are deemed to have violated PC, one of the most common claims about PC is that sanctions are often levied in some form of censorship (Hentoff, 1992; Hughes, 2010; Ford, 2017; censorship of others

and self-censorship will be discussed at length in a Chapters II, III, and IV, so it will only be touched on briefly here). By *censorship* of others, I mean taking action or supporting policy that restrains freedom of expression (Lambe, 2008). *Self-censorship* refers to avoiding expressing one's opinion or sharing information due to fear of social sanctions (see more in Chapter II). Some argue that PC ideologies often promote the censorship of dissenting views (Hughes, 2010; Ford, 2017; Loury, 1994; Robinson & Reid, 2016a). Though censorship has long been thought to be the domain of the right (e.g., conservatives often support the censorship of anti-government sentiments as 'unpatriotic'), liberal PC is also often associated with censorship of dissenting views (Hentoff, 1992). For example, in February 2017, violent protests broke out on the UC Berkley campus, leading administrators to "deplatform", or cancel an appearance by right-wing provocateur, Milo Yiannopoulos, that had been sponsored by Berkeley College Republicans. Though much of the violence and rioting was committed by non-students (Park & Lah, 2017), the university decided it could not provide enough protection for Yiannopoulos or those attending the event. In April, 2017, similar events occurred when UC Berkeley chose to cancel another speaking event by conservative pundit Ann Coulter because of safety threats. The cancelling of both Coulter's and Yiannopulos's events prompted accusations of PC culture causing the censorship of unpopular views (Peters & Fuller, 2017). Other well-known examples of deplatforming include Richard Dawkins and Bill Maher, who had speaking events cancelled at UC Berkeley because of their criticisms of Islam. Although both men have been publicly critical of religion in general and Christianity specifically, they were deplatformed after criticizing Islam, ostensibly because Muslims are a liberal PC protected group.

Liberals have been similarly censored due to threats of violence. For example, in 2017, when Princeton professor Keeanga-Yamahtta Taylor criticized Donald Trump for being a "racist, sexist, megalomanic" on Fox News, she was sent death threats and had to cancel her speaking tour. In these cases, universities certainly had good reason to cancel their events; they have an obligation to keep their students safe. But why is that those who protested in favor of cancelling such events (whether they resorted to violence or not) thought that these speakers should be censored? There are other ways that people can express their opposition to campus speakers without calling for deplatforming (e.g., peaceful counter protests, alternative events, engaging in debate). When social sanctions take the form of censorship, rather than these other methods of expressing dissent, it may decrease the number viewpoints to which students are exposed. Furthermore, censorship of others can send a message about what viewpoints are appropriate, and this has the potential to create a chilling effect and self-censorship (see Chapter II). If universities are places to invent, discuss, and debate ideas, then it is important to carefully weigh the costs and benefits of censoring speakers who might offend protected groups of students.

Defining PC

The key pieces of PC phenomenology include representations of social groups along the lines of protected classes and perpetrators, language to describe groups and actions, conflicting group norms, social influence, offense taking, and self-and other censorship. Given the importance of these components of the phenomenon, it is essential that a definition of PC allows for multifaceted phenomena. It will be worthwhile to consider previous attempts at defining PC before offering a formal definition. Historically, definitions of PC have been contested, and have frequently been assumed from popular culture rather than

from a scientific perspective. Further, among those who have offered scientific definitions, these definitions are often too narrow to allow for key components of PC phenomenology. Scholars have pointed to the difficulty of defining PC (Gauthier, 1997; Hughes, 2010; Simmons, 2008), arguing that all people have a slightly different interpretation of the term (Berman, 1992; Weigel, 2016). The reason for this ambiguity, according to these scholars, is that PC is not just one thing, but instead has been invoked to refer to many components of liberal politics (e.g., multicultural curricula, language use, speech codes, affirmative action) (Berman, 1992; Hughes, 2010; Robinson & Reid, 2016a). Because PC, as Hughes (2010) puts it, is "more easily recognized than defined" (p. 9), many scholars avoid defining the term altogether, and seem to hope that their readers will share their interpretations (e.g., Asante, 1992; Avery & Steingard, 2008; Van Boven, 2000). This lack of definition, however, has led to decades of scholars using the term PC despite studying different phenomena.

Evidence indicates that it is not just scholars who differ in their definitions of PC: the general population is similarly divided over what PC actually means. Lalonde, Doan, and Patterson (2000) conducted a study where they asked both college students and faculty how they defined the term "political correctness"[2]. The researchers grouped responses into three categories. First, 24% of respondents had a *literal* definition of PC, meaning they combined the words "political" and "correct", and interpreted them as meaning policies that are positive for society. A small group of participants (6%) defined PC in a way that the researchers coded as *right-wing backlash*, meaning participants defined "PC" as a pejorative term used by conservatives to attack liberals. The majority of participants (52%), however, had what

[2]Though this study was published in 2000, the data were collected in 1994 in Canada.

the authors refer to as a *popular* definition of PC, meaning avoiding offending marginalized groups. When repeating the study with faculty, the researchers found similar patterns. More recently, a content analysis of Tweets that include a form of the term *politically correct* found that people have begun to use the term in a more positive way (Wilkström, 2016). These studies show that not only do people vary in their definitions but that those definitions have changed over time.

Despite the difficulty of defining PC, some scholars have recognized the need for a definition. Those scholars who formally define PC do so in several of ways. Many scholars use the popular definition of PC, meaning avoiding speech that is derogatory toward marginalized groups (e.g., Avery & Steingard, 2008; Favreau, 1997; Gauthier, 1997; Hyde & Ruth, 2002; Maas, Suitner, & Merkel, 2013; Marques, 2009). Others claim to move beyond this popular definition, but the definitions they provide are still conceptually similar to the popular version. For example, Hughes (2010) argues that in order to be PC, it is necessary but not sufficient to avoid hate speech and racial slurs. One should also avoid language that could be viewed as marginalizing, prejudiced, and insulting towards protected groups**.** Though this "popular" definition of PC may be helpful in identifying the ways in which the public and politicians use PC, it is not broad enough to allow for other features of PC, such as attributions of offense (when and why are language, attitudes, and behavior are considered marginalizing towards these protected groups?) and social influence (when do interpretations of offense results in sanctions?).

Other scholars define PC not as a liberal phenomenon, but instead define the term in its pejorative sense: as a weapon conservatives use to downplay the legitimacy of arguments made on the left (e.g., Lea, 2009; Wilson, 1995). For example, Whitney and Wartella (1992)

point out that though the label "PC" has been used by conservatives as a way of attacking liberals, the movement originally "referred to a position – not readily attained – of liberal, progressive, and humanitarian concerns for the poor, homeless, disenfranchised groups of society" (pp. 85-86). Though these definitions shed some light on the PC movement, they also frame the debate simply as a competition between liberals and conservatives.

More comprehensive definitions acknowledge that PC is not just about particular norm content (e.g., norms prescribing the use of inclusive language), but instead involves intergroup competition between political adversaries. For example, Fairlough (2003) sees the PC movement "as a political controversy in which both those who are labelled as 'PC' and those who label them 'PC' are engaged in a politics that is focused upon representation, values, and identities – in short, a cultural politics" (p. 17). Similarly, Chait (2015) argues that PC "is a style of politics in which the more radical members of the left attempt to regulate the political discourse by defining opposing views as bigoted and illegitimate." These definitions broaden our understanding from looking at PC as a set of norms or beliefs about representation and discrimination, and instead allow for explanations of the ways in which these beliefs are advocated. While these definitions capture a great deal of the phenomenology of PC, it remains that these definitions still rely on the specific contexts in which PC occurs or is judged, and do not explicate the intergroup nature of PC, or its ideological elements (i.e., a particular vision of what society should be for all people and all times; see next section for an extended discussion of ideology).

Loury (1994) moves beyond defining PC in terms of a specific set of issues or intergroup competition, but focuses on freedom of expression. He defines PC as "as an implicit social convention of restraint on public expression, operating within a given

community" (p. 430). In defining PC this way, Loury avoids defining PC as purely the domain of the left, but instead allows room for any ideology to be classified as PC, given it implicitly encourages the inhibition of public expression of opinions. Similarly, Morris (2001) defines PC in the following way: "because certain statements will lead listeners to make adverse inferences about the type of speaker, speakers have an incentive to alter what they say to avoid that inference" (p. 233). Morris does not specify a particular topic or domain of speech content, but instead considers any realm in which this occurs to be an instance of political correctness. Though these definitions are useful, they only focus on one feature of PC: self-censorship. This is an important component of PC, but these definitions do not acknowledge the complexity of PC norms and its many of its ideological functions. Though Loury's argument uses examples in line with the popular definition of PC (e.g., examples about fear of offense stifling debate about university divestment from South Africa in the 1980s), the definition itself does not go far enough to include all of the features of PC as it does not explain why such social conventions would exist.

Formal Definition of PC

One of the main difficulties in defining PC is that because PC is multi-faceted (Berman, 1992; Hughes, 2010; Robinson & Reid, 2016a), some see the term as being derogatory and reject classifying communication or behaviors as PC (Lalonde et al., 2000), and others have a vague sense of what PC means and define it in terms of example. Consequently, many scholars write about PC but differ in their interpretation of communication and behaviors that are PC (Hughes, 2010). However, just because two different people might classify communication differently in terms of whether or not it is PC does not mean that scholars cannot or should not provide a scientific definition. Similarly,

people differ in the extent to which they interpret communication as racist or sexist, but that does not mean that scholars cannot or should not provide scientific definitions of *racism* or *sexism*. With that being said, one way to overcome the flaws of previous definitions of PC is to create a definition that describes the features of the phenomenon, but at the same time, does not specify the content of PC, but rather frames PC in terms of its communicative purpose.

In attempting to overcome the flaws in previous definitions and account for the various features of PC, Robinson and Reid (2016a) define PC as "a strategy in which political actors (on the left and right) use and promote ideologies that enable the control of discourse and by extension, political adversaries" (p. 87). There are several important components of this definition that allow it to describe the multi-faceted phenomenology of PC better than previous definitions. First, PC is a *strategy*, meaning it is functional and used to benefit those who are invoking it. This does not mean that it is necessarily used consciously. Research on motivated reasoning finds that rather than interpreting events and the behavior of others objectively, people draw the conclusions that are consistent with their desired outcome, and that this process occurs automatically (see reviews in Ditto et al., 2009; Kunda, 1990). Second, PC is enacted by *political actors*. These actors need not be those who are formally involved in politics, but are individuals who are promoting their interests. Relatedly, *political adversaries* refers to members of political groups or those who advocate opposing political ideas. Third, PC is used to promote *ideologies*, meaning shared interpretations about the way the world is and the way it should be (Jost, Fredirco, & Napier, 2009). Robinson and Reid note that PC is used on both the right and left, and this is consistent with Jost et al.'s (2009) classification of ideologies, which fall onto a continuum

that ranges from liberal to conservative. Last, PC is used to consciously or unconsciously control political adversaries through the control of *discourse*. The control of discourse refers to language, censorship, setting the terms for acceptable debate, and other norms regarding communication.

The definition put forth by Robinson and Reid (2016a) explicitly frames PC in terms of intergroup relations (e.g., political adversaries). Previous definitions have implied that PC is related to intergroup relations without explicitly referring to PC's intergroup functions. For example, popular definitions of PC describe it as regulating speech toward and about protected groups. Scholars who define PC in a pejorative sense view the term "PC" as a weapon in political conflict. Others, such as Chait (2015) and Fairclough (2003), define PC in terms of competition between political groups. Like Loury (1994), Morris (2001), and Fairclough (2003), this definition does not restrict PC to the domain of the left, or only involving norms related to the suppression of discrimination. Rather, by defining PC in terms of intergroup relations it is possible to allow for PC's features.

First, by defining PC as a strategy used by political actors against political adversaries, it is possible to see why PC must, necessarily, involve coalitional partners (i.e., protected groups), and perpetrators (i.e., political adversaries) (see next section for an extended discussion of victimhood). Second, by defining PC as a strategy, this definition better accounts for phenomena that include conflicting norms and attributions of offense. Conflicting norms make it easy to follow some PC prescriptions while inadvertently violating others, and this allows for political actors to strategically determine what is offensive (Robinson & Reid, 2015) based on what is consistent with their views on who is moral and who is immoral (Ditto et al., 2009). Third, by defining PC as the promotion of ideologies

(i.e., a vision of how the world should be; Jost et al., 2009), this definition can better account for a phenomenon that involves social influence and the social sanctioning of violators. Because ideologies represent a vision for the way the world should be, all people are thought to be equally subject to ideological prescriptions (Haidt, Rosenberg, & Hom, 2003), and therefore ingroup and outgroup members may both be subject to sanctions for violations. By framing PC as an ideology, then, Robinson and Reid's (2016a) definition allows for explanations of PC that include social influence and the social sanctioning of those deemed to have violated ideological prescriptions. Last, by identifying control of discourse as the mechanism through which PC ideologies are promoted, this can account for the normative prescriptions surrounding language, as well as why PC is often enforced through censorship and by encouraging self-censorship. Whereas previous definitions of PC have focused on a few of the features of PC, the definition put forth by Robinson and Reid (2016a) improves on previous definitions by being able to describe all of its characteristic features.

PC and Intergroup Relations

The definition of PC provided by Robinson and Reid (2016a), and conceptualization of PC that will be used in this dissertation, frames PC in terms of intergroup relations, rather than as a phenomenon that is specific to any one ideological perspective (i.e., not solely the domain of the left). Therefore, it is important to consider what function PC serves in intergroup relations and competition, and how this is distinct from other known intergroup phenomena.

Ideological Functions of PC

Robinson and Reid (2016a) argue that rather than PC being a norm, or even a cluster of norms, it is actually an ideological perspective that is advocated through the control of

discourse. Though most norms describe/prescribe behavior for the ingroup and do not apply to outgroups (Hogg & Reid, 2006), outgroup members are often compelled to adhere to PC (Ford, 2017; Robinson & Reid, 2016a). The ability of a group to develop a set of norms and then be able to punish outgroup members for violating the norms makes PC a powerful weapon in intergroup competition. Adherence to normative communication and behavior can be an important indicator of ideological commitment (e.g., the use of the most contemporary normative language). Norms, however, are just one part of ideological thinking. In their review on scholarship about ideology, Jost et al. (2009) conclude that though ideology is often not formally defined in research, at its most basic level, an ideology "helps to interpret the social world, and that it normatively specifies (or requires) good and proper ways of addressing life's problems" (p. 309). PC is not just a group of norms that describe the present, but is an ideology for how the world ought to be (Robinson & Reid, 2016a). Using Jost et al.'s (2009) definition of ideology, then, PC serves ideological functions.

Viewing PC as an ideology, rather than a norm or cluster of norms, is consistent with Tafjel's (1982) conceptualization of the three social functions of stereotypes. According to Tajfel, stereotypes about outgroups serve three functions. First, stereotypes can be used to justify "actions planned or committed against outgroups" (p. 22). For example, those on the right who support banning immigrants from predominantly Muslim countries from entering the United States can justify their actions by invoking stereotypes of Muslims as terrorists. Similarly, those on the left who want to control the discourse of those on the right can invoke stereotypes of Republicans as racists and misogynists. Second, stereotypes can be used as a simplified explanation of distressing social events. For example, people may invoke stereotypes of racist police officers to explain police shootings, just as unemployed US

citizens may rely on stereotypes of Mexican immigrants to explain the changing employment landscape. Third, stereotypes can be functional in that they promote positive group distinctiveness, such as liberals invoking stereotypes of conservatives as "dumb rednecks" or conservatives viewing liberals as "bleeding heart tree huggers." Tajfel's analysis of these social functions of stereotypes complements basic cognitive approaches (i.e., approaches to stereotypes that focus on how they enable individuals to reduce uncertainty about the social world) and understandings of stereotypes by placing them in a broader social and functional context. In the same fashion, liberal PC contains social norms regarding language use, multiculturalism, diversity, sexism, and so forth, as well as the functional benefit of controlling political adversaries.

According to Robinson and Reid (2016a), each of these three ideological functions of stereotypes are important to the formation of PC. If, as is argued in this dissertation, PC serves to further the interests (e.g., maintaining resources, protecting coalitional partners) of the ingroup by controlling rival outgroups through the control of discourse, then groups must justify the actions they take in order to gain that control. One way that this happens is through justifying the censorship of rivals by claiming that the outgroup has victimized the ingroup (See Chapter III, Study 1 for more on this topic). For example, during the McCarthy era, under the guise of national security, artists, journalists, and politicians were censored. McCarthy and his party were able to justify the censorship with the accusation that they were communists, and by invoking the stereotype that communists are dangerous and un-American (Hughes, 2010). Similar patterns can be seen amongst some of those who endorse liberal PC ideology. For example, Dredger (2015) details the story of J. Michael Bailey, a tenured professor of psychology at Northwestern University, who was almost fired for publishing a

book that claimed that one form of trans-genderism involves mental illness. His opponents even posted photos of Bailey's children online, justifying their actions by claiming he was a threat to the trans-community. This is all to say that PC, on both the right and left, is used as a tool to blame opponents for the state of the ingroup (whether that blame is legitimate or not) to justify actions against them.

PC and a Fourth Ideological Function

Tajfel's (1982) ideological functions of stereotypes are helpful in understanding PC. However, PC arguably serves a fourth function: the control of outgroups through the enforcement of social norms. Robinson and Reid's (2016a) definition of PC, which focuses on the control of adversaries specifically through the control of discourse, moves beyond Tajfel's functions of stereotypes by explaining how groups are able to leverage their own ideologies to engage in intergroup competition.

PC ideologies (both conservative, such as McCarthyism, and contemporary liberal PC), claim that their ideology is *moral* and non-subscribers are *immoral* (Ditto et al., 2009; Jost et al., 2009). *Morality* can be understood as "interlocking sets of values, practices, institutions, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make social life possible" (Haidt, 2008, p.70). Because PC is made up of broad interwoven norms and values, those deemed to have violated a PC norm are not sanctioned for a single norm violation, but they are accused of being immoral, as they have violated an entire view of how society should be (e.g., McCarthy claiming that reading "defamatory" literature was a danger to other people; leftwing PC advocates who claim that those who use non-inclusive language are discriminatory) (Ditto et al., 2009). McCarthy was able to enforce his beliefs on others not through arguing about differences in opinions about

what it means to be a Communist (or whether communism is harmful), but through accusing opponents of being disloyal and therefore immoral. Liberal PC similarly relies on morality, and with a particular focus on equality and caring for others (Haidt, 2012). Because the normative behaviors associated with PC ideologies fall along a moral dimension that subscribers believe should apply to all people, this allows them to police the behavior of outgroups.

Much of the public communication surrounding PC is characterized by accusations of immorality, which can be understood with reference to Hofstader's paranoid style (Robinson & Reid, 2016a). According to Hofstadter (1964), those who embody the paranoid style attribute conflict to simplistic good vs. evil explanations (i.e., moral vs. immoral). Those who embody the paranoid style do not see disagreements as differences in opinions or values, but instead attribute differences of political opinion to malicious intent and persecution (Hofstadter, 1964; Oliver & Wood, 2014). Importantly, Hofstadter (1964) notes that the paranoid style "has to do with the way in which ideas are believed and advocated rather than with the truth or falsity of their content" (p. 5). Describing the advocacy of contemporary PC in terms of the paranoid style is not meant to comment on the pervasiveness of discrimination in the US, just as describing McCarthy's advocacy in terms of the paranoid style is not a reflection on the pervasiveness of communism during his era. However, just as in McCarthy's era when disagreements between liberals and conservatives were met with accusations of disloyalty, liberal PC norms are often enforced in such a way that any disagreements, regardless of whether or not they are reasonable, are met with accusations of racism/sexism/heterosexism (Chait, 2015; Loury, 1994). For example, if an individual publicly expresses concerns about more open immigration policy, they may be met with

debate about the costs and benefits of such a policy. However, it is also possible they could be described as a racist and bigot by political opponents, simply for expressing a dissenting view. When advocacy of PC is characterized by the paranoid style, it may leave little room for dissent or for collaboration between sides (see more in section on Public Deliberation Chapters II & V).

When ideologies are framed in terms of morality, it makes it easier to justify coercing those who are perceived to be out of step with the ideology (ingroup and outgroup members alike) to act in ways consistent with it. Usually, normative attitudes and behavior are policed on an intragroup level. Group norms are useful for maintaining intragroup harmony by bonding the group together, and when group members do not abide by group norms, it dilutes the integrity of the group (e.g., Marques, Abrams, & Matrinez-Taboada, 1998). Therefore, liberals who violate liberal PC norms might be punished by fellow ingroup members (e.g., women being accused of being anti-feminist for choosing to be stay at home mothers; women who voted for Donald Trump being described as having "internalized patriarchy"). PC, on the other hand, is not just enforced on ingroup members but may also be enforced on outgroup members as well. Because PC ideologies are advocated in a way such that those who outwardly oppose the ideology are accused of immorality and can be sanctioned, PC can reflect and increase intergroup political competition and belief polarization.

When norms are enforced on both ingroup and outgroup members, as I have argued is the case for the normative prescriptions of PC ideologies, they have authority both over those who genuinely support them as well as those who do not. Turner's (2005) three process model of persuasion helps to explain how and why PC norms have such authority. According

to this model, people follow norms for one of three reasons. First, people who genuinely subscribe to PC, or "true believers," can be legitimately persuaded to follow a norm because they subscribe to the overarching ideology. These people adhere to PC norms (e.g., using PC language) because they genuinely think it is the right thing to do. In the absence of persuasion, people may also follow norms because a legitimate authority compels them to do so. For example, when the California Legislature passed bill No. 1887, which bans state money from being spent on employee travel to states with laws that discriminate against sexual minorities, they compelled state employees to abide by their ideology, regardless of whether or not they support the policy. Similarly, teachers who live in states that mandate abstinence-only sex education must abide by that policy, even if they think it is more effective to teach a broader sex education curriculum. Last, people may be coerced into abiding by PC norms due to fear of sanctions, such as job loss, social ostracism, or fines. Because issues of morality are supposed to be universal (Ditto et al., 2009), both ingroup and outgroup members should be expected to adhere to such codes of morality. When outgroup members express views that are inconsistent with PC, they may be accused of immorality (because only an immoral person would express such views), thereby giving the ingroup justification to engage in social sanctioning and censorship. In this way, PC serves the unique function of being able to be imposed on both ingroup members and outgroup members.

Conclusion

PC is multifaceted, and according to Robinson and Reid (2016a), its primary purpose is to control political adversaries through the control of discourse. There are ideologies on both the right and left that contain the features of PC (Robinson & Reid, 2016a). However, the decades-long debate about the costs of benefits of PC has largely been about liberal PC

and its effects on university life (i.e., using the popular definition) (e.g., Berman, 1992; Chait, 2015; Fairclough, 2003; Loury, 1994; Lukianoff & Haidt, 2015; Weigel, 2016). In order to contribute to this debate, this dissertation will test the relationship between PC, support for censorship, self-censorship, and public deliberation specifically within the domain of liberal PC. A question remains that has gone largely untested: is liberal PC related with self-censorship as some have suggested? If, as many have claimed, liberal PC has caused a chilling effect on college campuses where those who do not support liberal ideologies are afraid to speak out (Chait, 2015; Loury, 1994; Kipnis, 2015; Schlosser, 2015), then it essential to garner a better understanding of liberal PC and its role on college campuses. Self-censorship, its potential effects for public deliberation, and the rationale for the rest of this dissertation are discussed in Chapter II.

Chapter II: Political Correctness, Self-Censorship, and Public Deliberation

According to the definition used in this dissertation, I conceptualize PC as a tool used to control political rivals through the control discourse (Robinson & Reid, 2016a). While there are many ways to control discourse (e.g., by promoting some explanations for social events in favor of others), direct and indirect forms of censorship have captured most of the attention in the critiques of PC found in the humanities and social sciences (Berman, 1992; Hughes, 2010; Lukianoff & Haidt, 2015; Weigel, 2016). Though there are some features of PC that are unique to the political left (e.g., a focus on diversity, protection of historically disadvantaged groups), what is common to PC on both the right and left is that those who endorse these ideologies may promote them through restraining the expression of dissenters. Restraints on expression may take the form of censoring others or may occur through a chilling effect. Self-censorship is a practically and theoretically interesting form of social control. First, if self-censorship is sufficiently frequent, it may lead to a decrease in viewpoint-diversity (Lukianoff & Haidt, 2015), and this has the potential to lead to making decisions before exploring the full complexity of problems (Loury, 1994; Morris, 2001). Second, the promotion of views that are inconsistent with scientific evidence (Dreger, 2014) can both decrease scientific discovery, and inhibit technological innovation (e.g., disbelief in human causes of climate change on the right; belief in the medical dangers of genetically modified foods on the left). Third, it is also possible that self-censorship may undermine democracy, because it may prevent a full range of political opinions from being considered for debate, and because those who feel that they are silenced may become susceptible to populist, rather than rational, leadership (Loury, 1994; Hofstadter, 1965; Habermas, 1984).

However, although the argument that PC culture on college campuses leads to self-censorship has been made for over three decades (e.g., Berman, 1992; Chait, 2015; D'Souza, 1992; Hentoff, 1992; Lukianoff & Haidt, 2015; Robinson & Reid, 2016a), it has yet to be empirically tested. As such, several questions remain: Does PC create a chilling effect of dissent on college campuses? If so, under what conditions might the self-censorship of some people increase diversity of opinion by creating a more comfortable environment for those from historically disadvantaged groups? Under what conditions might self-censorship hinder public deliberation by discouraging opposing viewpoints and inhibiting decision making? How should universities balance these tradeoffs? In what follows, potential theoretical explanations for the connection between PC and self-censorship are discussed and hypotheses regarding self-censorship deduced. Drawing from social identity theory (SIT, Tajfel & Turner, 1979), Tajfel's (1982) work on the social functions of stereotypes, self-categorization theory (SCT: Turner, Hogg, Oaks et al., 1987), and research on public deliberation, I propose three broad hypotheses about the effects of PC on censorship of others, self-censorship, and public deliberation. These hypotheses will be tested across three studies. Detailed predictions that follow from these hypotheses will be presented prior to each study.

Conceptualizing Self-Censorship

Before delving into the potential consequences of self-censorship, it is important to note what self-censorship is, and more importantly, what it is not. According to Hayes, Glynn, and Shanahan (2005), there are many instances when one may avoid self-expression without self-censoring. Hayes et al. make distinctions between self-censorship and other forms of staying silent. They argue that *opinion expression inhibition* refers to any instance

when a person chooses not to share their opinion. *Self-censorship* is more specific, referring to instances of "withholding of one's true opinion from an audience perceived to disagree with that opinion" (Hayes et al., 2005, p. 299). Using this definition, it is possible to distinguish self-censorship from other communicative interactions that may involve keeping one's opinions silent because of personal shyness, because of a desire to avoid offending others, or to be polite. However, Hayes et al. (2005) are careful not to include the motives that a person may have for self-censoring in their definition. Using Hayes et al.'s definition, self-censorship can occur if a person fears social sanctions, wants to avoid an argument, wants to look like a good group member, and so on. According to this conceptualization of self-censorship, the only necessary conditions are that a person *perceives* the opinion climate to be hostile (meaning they think they are in the opinion-minority) and subsequently avoids sharing their opinion.

Where Hayes et al.'s (2005) definition is very broad and can refer to a wide variety of situations, Loury (1994), who was writing about the potential effects of PC on college campuses, gives a more specific definition. He defines self-censorship as instances when "members whose beliefs are sound but nevertheless differ from some aspect of communal wisdom are compelled by fear of ostracism to avoid the candid expression of their opinions" (p. 430). This builds on Hayes et al.'s definition in two important ways. First, as explained in Chapter I, a feature of PC is that social sanctions may be levied against those who violate PC, which has the potential to compel people to publicly adhere to the prescriptions of the ideology whether or not they actually endorse them. Loury's definition stipulates that in order for opinion expression inhibition to be classified as self-censorship, the motivation for such inhibition must include a fear of sanctions. A second important part of Loury's

definition is that it implies that there are negative consequences to self-censorship: that legitimate or "sound" opinions are left out of the conversation. This is important because, where Hayes and colleagues do not stipulate that self-censorship must have negative consequences (although it is implied in much of their research), Loury assumes that it does.

There are instances when self-censorship may produce an increase in diversity of opinion. For example, as is argued by those in support of PC ideology, it is possible that social norms that proscribe communication that might be deemed discriminatory may cause those with racist, sexist, or heterosexist views to self-censor, and this may result in an environment where those from marginalized groups feel more comfortable expressing their own perspectives (Firsh, 1992; Perry & Williams, 1992). According to muted group theory (Ardner, 1978; Orbe, 1998), because societal communicative norms are determined by the communication patterns of members of dominant groups, it can make it difficult for members of nondominant groups to use their own communicative styles, which may prevent them from entering deliberation at all. For example, Orbe (1998) interviewed people from nondominant groups (People of Color, women, LGBT, those from a lower socio-economic status) about how they communicatively navigate interactions with members from dominant groups. Commonly cited responses included engaging self-censorship and avoiding conflict. From this muted group perspective, then, if PC causes members of dominant groups to self-censor, this may allow more space for members of nondominant groups to share their perspectives, which has to potential to increase diversity of opinion.

Despite these potentially positive effects of self-censorship, it is also important to examine the potential negative implications. One instance of self-censorship is unlikely to result in dire consequences for campus life (Morris, 2001). However, if it is true that a

culture of self-censorship may decrease public deliberation, diversity of opinion, and campus life, at least some of the time (Morris, 2001; Loury, 1994; Lukianoff & Haidt, 2015), then it is important to consider when and why that happens, and how that might affect public discourse.

Theoretical Explanations for Self-Censorship

There are three main theoretical perspectives (Communication Competence, Communication Accommodation Theory, and Spiral of Silence) that have direct implications for our understanding self-censorship. In what follows, I outline each of these three perspectives, and highlight the areas in which they are useful, and limited, in aiding our understanding of the effects of PC on self-censorship. This discussion ultimately provides context for a broader intergroup approach, which I will argue is better suited to examining the wide range of phenomenology involved in PC, as well as the specifics of self-censorship.

Communication Competence

Where Loury (1994) assumes that self-censorship has negative consequences by definition, not all theorists agree that self-censorship is necessarily negative. For example, there may be times when self-censorship by one person may result in an overall improved communicative interaction. According to Spitzberg's (1991) model of intercultural communication competence, communication can be evaluated on two continua. First, communication varies in social appropriateness, or how closely it adheres to group and relational norms. Second, communication varies in effectiveness, or how well it accomplishes the goals of the communicator. Communication is considered to be competent when it is perceived as being high in both appropriateness and effectiveness. When communication is either appropriate or effective, but not both, according to this model, it is considered

suboptimal. For example, a person may cater to their audience, even at the risk of not expressing their true opinions (i.e., their communication is appropriate but ineffective; also called *sufficing*), or a person may deceive others in order to achieve their goal (i.e., their communication is inappropriate but effective; also called *maximizing*). According to this model, communication that is perceived as appropriate by receivers and is effective in accomplishing the speaker's goals (i.e., competent communication) is considered optimal. Spitzberg acknowledges there may be times that these "noncompetent" types of communication are required, but, he claims, "they are rarely the ideal behaviors to employ in any given circumstance" (Spitzberg, 1991, p. 380). Using this model, it is possible to conceive of self-censorship as a form of competent communication, depending on the speaker's goals. If a person's primary goal is to avoid social sanctions, and not sharing a dissenting opinion is also the socially appropriate thing to do (i.e., avoiding being perceived as offensive), then self-censorship can be competent communication, at least from the perspective of this model.

While this model may be helpful in explaining some ideals of effective communication in intercultural interactions, the goal of these interactions is often to have a positive social interaction. However, this model does not explain how receivers determine what is and is not appropriate (i.e., engaging in motivated reasoning when determining the appropriateness of a message), nor how speakers navigate the challenges of assessing what someone else will think is appropriate. By claiming that competent communication is optimal, it implies that appropriateness and effectiveness carry equal weight. However, there are times when it is not possible to be appropriate while still being effective. For example, someone may have a position that is factually correct but that opposes the dominant ideology

in the environment in which they are speaking. Such a person is likely to be considered "incompetent" by this model (i.e., minimzing, maximizing, sufficing) because those expressed opinions are counter to PC. On the other hand, a speaker could use illogical arguments that are deemed appropriate, successfully influence an audience, and thus be perceived as competent. Essentially, it is possible to be viewed as an incompetent communicator while expressing a logical position if it departs from a PC position, and it is also possible to be viewed as a competent communicator while expressing illogical but socially appropriate views. Public deliberation allows for (and requires) disagreement in order to settle on the best possible solutions to society's problems, which can be highly non-normative, and may be expressed in ways that are ineffective or viewed as inappropriate, but would nonetheless potentially offer useful perspectives to political debate.

Spiral of Silence

Noelle-Neumann's (1974) spiral of silence theory (SoS) was designed to address societal level consequences of self-censorship. She posits that people are more likely to self-censor when they think their opinion is in a minority, especially for important and controversial social issues (i.e., morally loaded topics). According to the theory, people have a quasi-statistical sense of what the majority opinion is, and when they perceive themselves as being in the minority, they will either express what they think is the majority opinion or they will avoid expressing their opinion at all. As fewer people express minority opinions, those opinions become less prevalent. Over time, more people silence what they *perceive* to be the minority opinion. Importantly, this process is assumed to apply only to politicized social issues, or cases where topics are "morally loaded" (Noelle-Neumann, 1974). Using the SoS, a prima facie case can be made that PC norms cause self-censorship. As I argued in

Chapter I, PC involves morally loaded issues about which there appears to be consensus on many college campuses (Chait, 2015). As journalists suggest, and dozens of cases demonstrate, PC norms are sometimes strictly enforced on college campuses (Weigel, 2016). The frequency and intensity with which PC violations are sanctioned (e.g., forced apologies, sensitivity training, suspensions) may indicate to students, faculty, and staff, that it is unacceptable to express opinions that counter PC ideology (Lukianoff & Haidt, 2015). Because interpretations of PC violations are subjective (Robinson & Reid, 2016a) even communication that is not intended to appear racist, sexist, homophobic, or derogatory toward any other PC protected class, can potentially still be subject to harsh social sanctions.

However, there are flaws in the SoS theory that prevent it from fully explaining how PC can lead to self-censorship. First is a theoretical flaw. According to Noelle-Neumann, there are "hard core" individuals who are willing to speak out even when they are in the minority. Despite the existence of these individuals, this does not appear sufficient to end the spiral of silence. Noelle-Neumann uses the findings of Asch's (1956) social influence studies as support for this argument. However, Asch (1956) actually found that if only one person in a group deviated from the norm, it reduced conformity to zero. Yet in the case of PC, there are many influential figures that speak out against PC norms (e.g., Donald Trump, conservative student groups, liberal academics, such as Jonathon Haidt), and still PC norms maintain authority, particularly on many college campuses. The SoS has trouble explaining why people who do not support PC norms would remain silent, and why the PC position maintains authority when dissent is not uncommon.

A second problem is that the SoS has received weak empirical support. In a meta-analysis of 17 published SoS studies, Glynn, Hayes, and Shanahan (1997) found a significant

but very small (r = .05) effect of perceptions of minority support and willingness to speak out. Hayes et al. (1997) attribute these small effects to the fact that SoS studies generally examine self-reported willingness to speak out rather than actual behavior, but it is not known whether this would resolve the lack of evidence. Another possibility, and one that is pursued in this dissertation, is that social norms are fluid, and context dependent (see Hogg & Reid, 2006). Despite Noelle-Neumann's argument that people constantly assess the opinion climate in order to avoid expressing minority opinions (the "quasi-statistical sense") (Noelle-Neumann, 1991), SoS researchers rarely test exactly how people assess this opinion climate (Hayes, Matthes, & Eveland, 2011), and often operationalize the opinion environment as a relatively static feature, rather like a snap-shot opinion poll. Rather than considering when and how norms become socially relevant for behavior, researchers typically ask participants to imagine that their opinion is either in the minority or in the majority (e.g., Hayes, 2007), or how popular they think their opinions are (e.g., Matthes, Morrison, & Schemer, 2010) (see Glynn et al., 1992 for meta-analysis). This methodological concern is particularly relevant to the study of PC, because, as was described in Chapter I, PC norms often conflict with each other which can create uncertainty about the opinion climate. People are very rarely given explicit information about support for PC, but instead rely on social cues to assess the opinion climate. It is important to examine how people evaluate social cues to determine what the opinion climate is and whether they should express themselves. Overall, the flaws in SoS prevent it from being able to explain how PC affects opinion expression and self-censorship.

Communication Accommodation Theory

Communication Accommodation Theory (CAT: see Dragojevic, Gasiorek, & Giles, 2016 for review) can better explain how people may make the decision to self-censor, and what subtle forms that self-censorship might take. CAT explains how people adjust their communication to manage social distance and communicative goals in relation to communication that can take place at interpersonal and/or intergroup levels. Though the theory was originally used to explain style shifting in accents, dialects, and languages, it has also been applied to other areas of communication, including nonverbal communication and topic avoidance (see review in Soliz & Giles, 2014). People converge toward their communication partner (i.e., adjust their communication style to be closer to that of their partner) to show affiliation, and diverge from their communication partner (i.e., adjust their style away from their partner) to maintain or amplify social distance. Though there are times when people converge toward their communication partner's idiosyncratic communication style, these communicative adjustments are often based on stereotyped expectations of a person's communicative style (Dragojevic et al., 2016). This adjustment can occur symmetrically, meaning both partners converge toward each other, or may be asymmetrical, meaning only one communication partner converges toward the other.

According to CAT, if a person who holds an opinion that they know/expect is discordant from their communication partner, and they want to be evaluated positively by that partner, they will likely accommodate, and self-censorship is a potential form of accommodation. Furthermore, Robinson and Reid (2016a) argue that PC ideologies have broad reaching social influence, such that even those who do not endorse the ideology are still expected to follow its normative prescriptions. From a CAT perspective, one way in which this might manifest itself is through asymmetrical accommodation: when

communicating about topics in which PC is salient, a person who does not support PC may be more likely to converge toward their perception of what is expected by a PC supporter than vice versa. In this way, CAT is a useful theory for examining how people accommodate to normative PC language as well as make the decision to discuss or avoid PC relevant topics.

CAT is useful in explaining why people may accommodate toward their communication partner (be it through converging toward their language style, expressing a normative attitude, or avoiding a subject that might result in the perception of divergence). In addition, it is also useful in understanding why convergence about PC topics may generally go in one direction (convergence toward PC normative communication). Importantly, however, CAT is part of a general social identity perspective—which includes ethnolinguistic vitality theory, social identity and self-categorization theories. Combined with these theories, CAT generates predictions regarding the wider social implications of self-censorship, as well as predictions about who would self-censor, when, and with what consequences.

Asymmetric convergence has been examined within the context of languages. Research on ethnolinguistic vitality (Giles, Bourhis, & Taylor, 1977; Giles & Johnson, 1981; see review in Yagmur & Ehala, 2011) has used the predictions of CAT in order to examine why languages thrive or die. According the theory, the status of the language and its speakers, demography (the number of people who speak the language), and the extent to which the language receives institutional support (e.g., through being taught in schools or made an official language of a country), determines the language's overall vitality. When people feel their ethnolinguistic group is of low vitality, they will start to diverge from their minority languages and converge toward the dominant language. Over time, this consistent

and widespread asymmetric convergence toward the dominant language leads to the "death" of the minority language. Using the CAT perspective on ethnolinguistic vitality, it also possible to imagine how PC could lead to the loss of diversity of opinion: PC often appears to have broad support from students and faculty (i.e., demography), and often has wide institutional support (e.g., campus safe spaces, classroom climate clauses in syllabi). Though I am not aware of CAT research that examines diversity of opinion, if the consistent and widespread asymmetrical convergence can lead to the loss of diversity of languages, then the asymmetric convergence toward one ideology should also lead to the loss of diversity of ideas on college campuses.

Intergroup Relations, PC, Self-Censorship, and Public Deliberation

PC is an inherently intergroup phenomenon. By definition, PC is a tool used to engage in intergroup competition. Therefore, the communication competence model, SoS, and CAT can partially, but not fully, explain the relationship between PC and self-censorship, and what societal level effects that might have. Broader theories of intergroup relations, namely SIT and SCT, can offer a more complete explanation.

Before building hypotheses grounded in intergroup relations, it is important to explicate the different social contexts that I will examine in this dissertation, and explain why an intergroup approach is best suited for examining these contexts. This dissertation will examine three different contexts, all of which are relevant to university life. Contemporary PC has long been associated with universities and college campuses. Because this is the first empirical test of the relationship between PC and self-censorship, an examination of the phenomenon within the university context has the potential for important practical implications. Study 1 will examine reactions to guest speakers on campus who espouse

views that may be seen as offensive for people with different political views. Study 2 examines students' self-censorship in a hypothetical classroom scenario. Study 3 examines how participants communicate with a peer (a research confederate) about the topic of allowing speakers to present on campus, even if they might offend certain students. Taken together, examining these three different contexts can give a more complete picture about PC, censorship, self-censorship, and deliberation on college campuses (and UCSB in particular).

Although each of these contexts is different, they are all contexts in which group membership is likely to be salient, thereby making group memberships and social identities relevant (see next sections for a more in depth discussion of social identity and self-categorization theories). Dragojevic and Giles (2014) propose that interactions can range along two orthogonal continua: and *interpersonal* continuum (meaning interactants are focused on each other's unique traits and personality), and an *intergroup* continuum (meaning the extent to which interactants are focused on each other's group membership). Interactions that are highly interpersonal but low in intergroup might include two old friends reminiscing about memories, whereas an interaction that is highly intergroup but not interpersonal might include soldiers from different countries on the front lines of a battlefield. However, many human interactions make both personal and social identities salient. Consistent with the argument put forward by Robinson and Reid (2016a), as well as with the assumption made by other writers (e.g., Fairclough, 2003; Loury, 1994), I argued in Chapter I that PC is inherently an intergroup phenomenon. In what follows, I will make the case that PC necessarily makes group membership salient, even in interpersonal interactions, thereby making discussions about PC relevant topics intergroup. So, for example, even though Study

2 examines the context of a classroom discussion, and Study 3 examines an interpersonal conversation, because both of these contexts still involve PC norms, so group membership should still be salient. Therefore, an intergroup relations approach is well suited for hypothesizing about the connection between PC and the communication outcome variables of interest (support for censorship, self-censorship, public deliberation).

Social Identity Theory

Social identity theory (SIT: Tajfel & Turner, 1979) can provide insight into the intergroup relations that affect self-censorship. According to SIT, people derive part of their self-concept from social groups, and want to feel positively about the groups to which they belong (i.e. maintain a positive social identity). If group members do not feel positively about their group memberships (e.g., because the group is low status, because the group is accused of immorality), they can engage in three methods for changing their status. First, people can engage in social mobility, wherein they disassociate from the previous group memberships in order to join a new group (e.g., after Donald Trump failed to single out and denounce violent White Supremacist protesters in 2017, many business leaders stepped off his business council in order to avoid being associated with the group). When social mobility is undesirable or not possible (in the case of unchangeable group memberships), group members can engage in social creativity, wherein they try to create a positive image for the ingroup, without engaging in direct confrontation with an outgroup (e.g., when those involved in the body positivity movement say "big is beautiful") or the dimensions on which they are compared to outgroups (e.g., the Health at Every Size movement reframes obesity to be about health rather than weight). Last, when these strategies are not possible, groups can engage in social competition (e.g., war, violence, political battles, and censoring rival groups, such as is

happening with the free speech of right wing extremist groups following the Charlottesville protests this year). Social competition comes about when groups compete over resources (material or otherwise). Because these resources are scarce, by definition, there must be winners and losers (i.e., those who win/maintain the resource, and those who lose it) (Tajfel, 1982; Turner, 1975).

SIT predicts that when people move towards a social competition orientation and away from social mobility or social creativity orientations, they become more highly committed to their ingroups, and this should result in greater political polarization. For example, when far left and far right groups begin to publicly endorse and justify political violence, moderates who show support for the speech or actions of the outgroup may be sanctioned. At the same time, both sides will attempt to promote outgroup stereotypes that justify actions against the outgroup—such as labeling supporters of Black Lives Matter as terrorists, or labeling supporters of President Trump as fascists. In fact, politics in the United States are more polarized than ever, both in terms of politicians' voting patterns, which can be predicted by group membership alone (Andris, Lee, Hamilton et al., 2015), and citizens' social networks, which have become increasingly politically homogenous (Bahksy, Messing, & Adamic, 2015). Though engaging in social competition often involves violent intergroup conflict, it can result in social change. For example, when groups that have historically held little political influence compete with those who hold power, there is the possibility of shifts in power (Tajfel & Turner, 1979).

While intergroup competition can promote social change, it can also produce a range of other effects, both within and between groups. Robinson and Reid (2016a) argue that groups are more likely to rely on PC ideologies during times of increased intergroup conflict,

as that is when groups have the greatest incentives to maintain order within the group by sanctioning dissent, as well as censor rival outgroups (cf. Tajfel, 1981). This can be achieved through explicit censorship (e.g., not allowing dissenters to hold public presentations) or sanctioning those who have publicly expressed dissenting opinions. However, when one group is publicly censored (or publicly punished for dissent) by outgroup members, other group members are likely to take notice. For those who wish to avoid the risks of engaging in direct social competition, they may decide to self-censor rather than risk social sanctions. This chilling effect can be an effective way of engaging in social competition because it forces moderates to pick sides or politically disengage. SIT provides a useful perspective for understanding how and why groups can use the censorship of rival outgroups in order to create a chilling effect, thereby further promoting the ingroup's interests.

H1. Intergroup competition and support for censorship. Tajfel (1982) proposed that groups use stereotypes in order to explain their current position in society, especially under threatening conditions. When groups are engaged in social competition, they tend to rely on perceptions that they have been wronged (i.e., victimized) by the outgroup in order to justify actions against that outgroup. Research on competitive victimhood (CV) has found that groups that are engaged in violent conflict often make claims about being victimized in order to justify actions against their rivals (Noor, Schnabel, Halabi, & Nadler, 2012; an expanded review of competitive victimhood literature can be found in Chapter III). Though this research has largely examined how groups justify violence in particular, another possibility is that groups may use claims of discrimination and victimhood in order to justify censorship and the control of discourse. For example, many of the speech codes on college campuses in the 1990s, which regulated speech related to race, gender, and sexual

orientation, were justified by the claim that these groups had been the victims of historical injustices, and therefore needed to be protected from speech they might find offensive (Hentoff, 1992; Hughes, 2010). Similar calls of victimhood occur on the right (albeit, less often in the university context). For example, far right groups (e.g,. white nationalists, the "alt-right", and KKK) frequently use claims that there is discrimination against Whites in America to justify their violent actions, as well as hostility (typically online) towards members of groups such as Black Lives Matter and feminists. Referring back to the aforementioned contexts, when examining students' attitudes towards bringing potentially offensive guest speakers to campus, how might group membership, and in particular, political partisanship, be related to attitudes towards censorship? Based on SIT, Tajfel's (1982) ideological functions of stereotypes, and Robinson and Reid's (2016a) argument for how this relates to PC, I propose the following hypothesis (to be tested in Study 1, Chapter III).

> H1: Perceptions of victimhood will diverge across political lines, will be amplified by ideological partisanship, and the extent to which groups are perceived to be the victims will be correlated with the desire to censor those who have victimized them.

Self-Categorization Theory

Self-categorization theory (SCT: Turner et al.,1987) builds on SIT by specifying the cognitive processes that lead any given social identity to become psychologically operative (i.e, salient). SCT posits that people have identities that operate at different levels of abstraction: the human level (i.e., seeing the self a member of humankind), the group level (i.e., social identities, such as gender, nationality and political affiliation), and the personal

level (i.e., based on idiosyncratic characteristics that define individuals as distinct from others). Depending on the social context, different levels of identity are made salient. For example, two people discussing sports will find that their personal identities become salient when they share a sports group membership, but have differing opinions about, say, the performance of different athletes in their team. Personal identities become salient because they enable people to keep track of personal differences within intragroup social contexts. However, if those same individuals were to discover that they are fans of different teams, then their social identities would become salient, particularly if their teams are currently competing and/or have a historical rivalry. Social identities become salient in intergroup contexts, and enable people to keep track of coalitions. But more than this, when groups have intense intergroup conflict, social identities do not merely become cognitively operative, they also harden ingroup commitments, and have evaluative and emotional significance.

When social identities become salient, people psychologically represent their social identities as prototypes, or fuzzy sets of features that maximally clarify the pattern of ingroup similarities and intergroup differences. This may take the form of traits (e.g., skin color, group differences in average height), normative attitudes (e.g., political opinions, music preferences, support of sports teams), and normative behavior (e.g., dress codes, speech styles). Cognitively, group prototypes help people make sense of their social worlds, even if a person's behavior or communication does not give fool-proof cues to their group memberships (Hogg & Reid, 2006). For example, in a discussion of immigration, if a speaker who has not explicitly stated their political position uses the term "illegal immigrant," listeners are likely to categorize that person as conservative and in favor of stricter immigration policy, as that person is using normative language for conservatives (i.e.,

a prototypical behavior). Likewise, if a person used the term "undocumented," listeners are likely to infer that the person is liberal and in support of more open borders. People use the normative fit between political groups and language to make inferences about who is and who is not "one of us".

SCT also explains when and how self- and other censorship operates. When intergroup competition increases, prototypes polarize—intergroup differences between group prototypes are amplified, and the range of positions that fit the ingroup prototype decrease. Further, when intergroup competition is intense, prototypes do not merely describe patterns of individual variation, but they can also become narrow, prescriptive norms (see Hogg & Reid, 2006). When this happens, being a 'good' group member is to be a prototypical group member. When prototypes are highly polarized, people do not tolerate outgroup positions, particularly prototypical outgroup positions, and they also have low tolerance for deviance from the ingroup prototype that even implies slight support for the outgroup. For example, Senator Diane Feinstein (a Democratic senator) recently faced intense criticism from within her own party for stating, with many provisos, that President Trump "can be a good president," if he learns from his mistakes and changes. California State Senator De Léon claimed that Feinstein's comment "wasn't the proper tone or tenor, especially given the current state of politics at the national level," (Mehta, 2017), indicating that he thought that national politics were polarized enough that any discussion of compromise was inappropriate. When intergroup competition is intense, and prototypes become narrow and highly prescriptive, deviants face social sanctions, thus incentivizing non-deviance. In the case of politics and self-censorship, the implications are clear. If you hold a position that is not entirely consistent with the PC norm, it may be best to avoid expressing that opinion.

Hypothesis 2: Ideological primes and self-censorship. SCT predicts that people will use traits and behaviors to categorize themselves and others. Given this, people should use language to make estimates of political affiliations, and if that language makes PC norms salient, people should be more likely to self-censor to avoid potential sanctions if they have views that can be interpreted as discordant with those norms (see Ford, 2017; Robinson & Reid, 2016a). SCT is in agreement with Loury (1994) that "suspicious speech signals deviance because once the practice of punishing those who express certain ideas is well established, the only ones who risk ostracism by speaking recklessly are those who place so little value on sharing our community that they must be presumed not to share our dearest common values" (p. 437). If this is the case, in the context of teacher-student interactions in a college classroom, then we should expect that a teacher's use of language gives clues about the group norms and prototypical positions to students. Those who value the classroom and campus community (i.e., want to be prototypical UCSB students) should be less likely to share their opinions if they are likely to be perceived as deviant from the salient PC norm. Therefore, I propose the following hypothesis that will be tested in Study 2 (Chapter III).

> H2: PC language primes PC ideology, and is a cue to those with dissenting opinions that they should self-censor.

Why Self-Censorship Matters: Public Deliberation and Democracy

In this dissertation, I explore the relationship between PC and self-censorship. If, as was proposed in H2 and will be tested in Study 2, PC does cause self-censorship among certain groups, then it is possible that PC may be harmful to public deliberation and democracy. This idea that decreasing public deliberation is harmful to democracy is

axiomatic (see review in Carpini, Cook, & Jacobs, 2004), and is not tested in this dissertation.

According to theories of deliberative democracy, democracy occurs when all citizens feel that their interests are represented, and even when they disagree with policies, they will continue to see the system as legitimate (Carpini et al., 2004; Chambers, 2003; Dahlberg, 2005). In order to have a system of deliberative democracy, by definition there must be public deliberation (Carpini et al., 2004; Mendelberg, 2002). There are many views of what this "ideal" of deliberation might look like. Habermas (1984; 1989), for example, posits setting ground rules for engaging in debate, claiming that debaters should aim to use coherent arguments, be truthful and give reasons for their arguments, and that anyone who desires should be included in the conversation. Public deliberation not only allows for people to hear various viewpoints and consider alternative opinions, but this deliberation also encourages participants to reconsider their previously held beliefs in light of new arguments and evidence (Jackman & Sniderman, 2006).

The necessary conditions for deliberative democracy have become a popular area of scholarship in political science. According to Carpini et al. (2005), public deliberation does not include discussions between elites, but instead occurs in communication between citizens, including formalized discussions between citizens (such as at a town meeting), informal political talk, which can be face to face or mediated, discussions at public forums, and can be about any topic of civic importance. Chambers (2003) describes public deliberation as a scenario in which "debate and discussion is aimed at producing reasonable, well-informed opinion in which participants are willing to revise preferences in light of discussion, in information, and claims made by fellow participants" (p. 309). If, as Chambers claims,

democratic deliberation requires that participants be open to the possibility of being persuaded by each other's arguments, then it is possible to see how PC can be harmful to this goal.

Though widely researched and widely endorsed, there are some critics of this view of the public sphere. A small contingent accuses deliberative democracy scholars, and particularly Habermas, of being idealistic (for review, see Carpini et al., 2005). For example, Dahlberg (2005) argues that not all people are trained in rational argumentation, those without this training are left out of deliberation, and that those who come from non-dominant cultures may have different communicative patterns, and therefore may deliberate in equally legitimate, albeit different, ways. Though this claim of Dahlberg is yet to be empirically tested, if his claims have support, then this even further highlights the importance of allowing students to engage in deliberation in the classroom. If public deliberation is necessary for democracy, and most citizens are underprepared to engage in such deliberation (e.g., build coherent arguments) (Chambers, 2003), then it would seem that university educators should have a vested interest in preparing students to be effective deliberators, not only by allowing them to express their ideas, but also by exposing them to ideas that they may find upsetting and teaching them to engage with and challenge those ideas.

The empirical evidence surrounding what conditions are necessary in order for deliberation to take place, has been mixed. Reviews on the subject (see Carpini et al., 2005; Mendelberg, 2002) conclude that the evidence of the effectiveness of public deliberation is mixed. It depends on who is deliberating and how effectively they follow the rules of true deliberation (e.g., being open to other viewpoints). For example, research has found that face to face communication can foster a desire to cooperate (e.g., Orbell, Van der Kragt, &

Dawes, 1988; Kerr & Kaufman, 1994), but it can also increase the salience of group difference when group interests are at odds with each other (e.g., Bornstein, 1992; Bettencourt & Dorr, 1998). Similarly, some research has found that the opinions of the majority tend to be dominant in deliberations because groups with more members are able to produce a higher quantity of novel arguments (Mendelberg, 2002; Vinokur & Bernstein, 1978). However, there are also times when minorities can encourage majorities to consider alternative perspectives (Martin, Martin, Smith et al., 2007; Moscovici, 1980; Moscovici & Personnaz, 1980). Other research had found that those who are part of powerful groups are given disproportionate time in debates, leaving out those who come from disadvantaged groups (Mendelberg & Oleske, 2000). Though this work may provide mixed findings as to when rational deliberation is likely to occur, none of this undermines the main tenant of deliberative theory. A higher quantity of rational arguments present in the public sphere benefits public deliberation, makes people feel more efficacious about the democratic process, and is better for democracy (see Carpini et al., 2005). If it is stipulated that more rational deliberation is better for democracy, then it is important to consider the conditions under which an individual is more likely to engage in argumentation that exhibits the necessary conditions of rational deliberation.

Integrative complexity. Research on integrative complexity (IC: Tetlock, 1983) is better able to explain when a why a person is likely to engage in the cognitive processing necessary for public deliberation to occur. The ideals of rational deliberation are embodied in the concept of IC, which refers to the extent to which a person differentiates between different sides of an issue, which implies that the person recognizes different viewpoints as worthy of consideration. Integration requires differentiation, and refers to the extent to which

a person weighs different viewpoints to develop solutions that recognize or transcend those perspectives. People who think in low differentiated ways tend to frame arguments as good versus bad (cf. Hofstadter's paranoid style), whereas those who are high in differentiation recognize differences of opinion as worthy of consideration. Higher integrative complexity is more cognitively demanding, and indicates that the individual is attempting to reconcile positions. Past research has found that liberals tend to exhibit higher IC than conservatives (Brundidge, Reid, Choi, & Muddiman, 2014; Golec de Zavala, Cisak, & Wesolowska, 2010; Tetlock, 1983; 1985), as conservatives tend to have a lower tolerance for uncertainty (Jost, Glaser, Kruglanski, & Sulloway, 2003). However, not all research has found this to be the case. Conway, Gornick, Houck et al. (2016) note that there are domains in which people from any ideological perspective are likely to exhibit lower IC, but much of the empirical research on IC has examined domains about which conservatives are more likely to exhibit lower IC.

Independent of political orientation, people tend to exhibit lower IC about issues to which they have strong ideological commitments. For example, Tettlock (1986) found that people who equally weight a variety of values when making decisions about policy preferences are more likely to exhibit greater IC. For example, a person's policy preferences on whether or not taxes should be raised in order to fund programs for the poor could rest on several different values (e.g., social equality, individual mobility). For some people, their ideological perspective may greatly outweigh one of these values, and these people are likely to exhibit lower IC when discussing the issue. People who more equally weight these values tend to exhibit higher IC when discussing the topic, as they are more likely to recognize differences of opinions as worthy of consideration and integrate these viewpoints into a

solution that satisfies multiple values. Similarly, Conway et al. (2016) has found that people on both the right and left tend to exhibit lower integrative complexity when discussing issues to which they are ideologically committed when those ideologies are made salient, ostensibly because ideologues often see little room for nuance. For example, Conway et al.'s findings indicate that conservatives tend to be more dogmatic, less complex thinkers about religion and abortion, but liberals exhibit this same pattern when thinking about environmental protection and the death penalty. The researchers argue that this is because for those who are ideologically committed, when discussing issues that are important to their ideology, rather than simply seeing them as their own political positions, they think "all people who disagree with me are fools" (Conway et al., 2016, p. 783). This type of thinking is ultimately harmful to the democratic process as it decreases the chances that those from opposing ideologies will collaborate to find solutions for social problems.

H3: PC and Public Deliberation. PC ideologies often include a dogmatic style of thinking, as PC may provide simple explanations of right and wrong (Robinson & Reid, 2016a). According to SIT and SCT, when groups are in conflict, people should be more likely to rely on their social identities when creating explanations of their social worlds, and should be more likely to advertise their prototypically by expressing attitudes normative to the ingroup. Considering that people use these interpretations of their social worlds in order to justify actions against the outgroup (Tajfel, 1982), these explanations should be less complex and more dogmatic as it is easier to justify engaging in conflict when there are clear distinctions between right and wrong. PC ideologies reflect these simplistic explanations for social problems. If, as was proposed in H2, that language can prime PC ideologies, then those who endorse the ideology should make less integratively complex arguments when it is

primed, meaning they are less likely to engage in rational deliberation. I propose a third broad hypothesis to be tested in Study 3 (Chapter IV):

> H3: Those who support PC will exhibit less rational deliberation (as measured by integrative complexity) in debate about a PC relevant topic than those who do not support PC, particularly when primed with PC language.

Chapter III: PC and Perceptions of Group Victimization

According to Tajfel (1982), groups use stereotypes to explain social events, promote ingroup status, and justify actions committed or planned against outgroups. Stereotypes typically satisfy all three functions simultaneously. For example, a stereotype that describes an outgroup as "aggressive criminals" produces a comparatively favorable stereotype for the ingroup ("we are good, they are bad"), the stereotype offers a simplistic explanation for conflict ("they are bad people who want to steal our resources"), and also justifies actions, such as engaging in preemptive attacks, because doing so can be framed as a legitimate form of self-protection. Robinson and Reid (2016a) extended this analysis to the social functions found in contemporary PC, which contains a number of stereotypes regarding group victimhood via discrimination (i.e., victimhood). If PC contains these social functions, then it may be that perceptions of group victimization (e.g., through sexism and racism) are used to justify actions against outgroups (e.g., censoring outgroup speech, explaining and justifying violence) as well as promoting ingroup interests (e.g., endorsement of political policies). For example, those who promote liberal PC often support limits on hate speech because they believe that such censorship is necessary to protect members of marginalized groups. Those on the right also use accusations of victimization in order to justify actions against outgroups[3]. For example in a recent national survey, 27% of Republicans reported that

[3]Discrimination can be tracked empirically, and groups' claims regarding who receives the most discrimination could, in principle, be settled. However, the point here is that there are psychological and social reasons that people have perceptions of group victimhood that are at least independent of these empirical facts (Noor et al., 2017). The purpose of this chapter is not to argue that all groups who claim victimhood or discrimination are equally worthy of social recognition of their victimhood status. Instead, the purpose is to show that the *perception* of victimization is related to groups' engagement in social competition. In other words, the focus is on the psychological and communicative aspects of victimhood, rather than their ethics or politics.

Blacks were the victims of discrimination, and 43% of Republicans reported that Whites are actually the victims of discrimination (Jones, 2017), and this perception of White victimhood has been used as justification for many white supremacy groups (Bacon, 2017). Though Robinson and Reid (2016a) argue that perceived victimhood status is a fundamental component of PC, this question has not yet been tested in relationship to groups' attempts to control discourse. If Robinson and Reid are correct that victimhood status is a part of a functional ideology that serves to justify competition with members of political outgroups, then we should find that greater perceptions of victimhood (in the form of discrimination) are associated with greater support for censoring outgroup speech. Further, if the effect is a result of coalitional behavior driven by social identity processes, people who are more politically aligned with the left or right should be more likely to perceive ingroup victimhood than moderates. Thus, the dual purposes of Study 1 are to test whether liberals and conservatives are divided on the question of group victimhood, and whether this, in turn, predicts their support for censoring political rivals based on their political partisanship. This will be examined within the context of students' support for banning potentially offensive guest speakers from campus.

Competitive Victimhood

Research on Competitive Victimhood (CV, e.g., Iyer, Jetten, Branscombe et al., 2014; Jetten, Schmitt, & Branscombe, 2012; Schnabel & Noor, 2012; Sullivan, Landau, Branscombe, & Rothschild, 2012) shows that groups engaged in conflict often justify actions against outgroups by claiming that they have been victimized either more than others by their adversaries, or more unjustly. Further, CV increases ingroup cohesion, delegitimizes support for reparations for past wrongdoings, and garners support from coalitional partners (for

review, see Noor, Shnabel, Halabi, & Nadler, 2012). Engaging in CV can also serve a psychological benefit and help group members maintain a positive social identity, by allowing groups members to think that they are more moral than their rivals, thereby allowing them to engage in downward social comparisons (Branscombe, 1998; Noor et al., 2012). Ultimately, these psychological benefits of CV serve to give groups justification for actions against rival outgroups (Sullivan, Landau, Branscombe et al., 2012). In particular, groups often use CV has a justification for committing violence against rival outgroups (see reviews in Noor et al., 2012, 2017). Groups that engage in CV are less likely to reconcile (Noor et al., 2012) and are more likely to support extremism and fundamentalism (Young & Sullivan, 2016). Overall, the bolstering of victimhood status and the rejection of blame can have detrimental effects on reconciliation.

Though competitive victimhood generally refers to relations between groups engaged in intractable violent conflict, such as war (e.g., Hutus and Tutsis, Israelis and Palestinians, groups in Northern Ireland; see review in Noor et al., 2012), similar process occur in nonviolent (or less violent) intergroup relations (Young & Sullivan, 2016). Furthermore, it is not required for a group be recognized by other groups or governments as a victim group in order to engage in competitive victimhood. For example, advantaged groups may claim victimhood in order to prevent the redistribution of resources (Noor et al., 2012) (e.g., Whites' claims that affirmative action is tantamount to "reverse discrimination"). Sullivan et al. (2012) found that when men were primed to think about men's discrimination against women, they were more likely to claim that men are also the victims of discrimination. Similarly, Phillips and Lowery (2015) found that when Whites were primed to think about White privilege, they were more likely to report that they had been the victim of a personal

hardship, ostensibly to justify that the individual had earned their position in life, and to justify opposition to institutional changes that aid marginalized groups.

CV research suggests that political groups are likely to develop beliefs regarding competitive victimhood, and do so as a function of their political identification and ideology. Specifically, people on the political left and right should be more likely to perceive that members of their ingroup are victims of the actions of their political outgroup. But more than this, political groups on the left and right are typically based on coalitions—it is not a simple matter of "us" versus "them." On the political left, there are seven groups that are commonly referred to as having been disadvantaged by racism, sexism, homophobia, and colonialism (respectively, Blacks, Latinos, women, non-heterosexuals, Muslims, and Native Americans; Leo & Haidt, 2016). On the political right, coalitional members appear to be mostly Christians, men, and White people (Bacon, 2017). If CV does take place politically, these respective groups should be perceived as victims of the political outgroup:

H1a: The more liberal a participant is, the greater victimization they will perceive among racial minorities, sexual minorities, religious minorities, and women

H1b: The more conservative a participant is, the greater victimization they will perceive among Whites, Christians, and men.

Groups may claim victimhood in order to maintain a positive social identity by engaging in downward comparisons with their perpetrators (Noor et al., 2012). They also claim victimhood in order to justify engaging in conflict with outgroups (e.g., Noor et al., 2012; Tajfel, 1982). Though this victimhood is often used to justify violence, actions that groups take against rivals can take on many forms, including restricted access to material resources and maintaining structural inequality (Noor et al., 2012). Another possibility, and

one that has not yet been explored in research on CV or more generally within intergroup communication, is that groups may potentially use claims of victimhood in order to justify censorship of adversaries:

H2a: The more liberal a participant is, the more they will support banning conservative speakers.

H2b: The more conservative a participant is, the more they will support banning liberal speakers.

Finally, if political competition is, in essence, competition between members of groups with valued social identities, then people who have stronger political identification should be more likely to both endorse CV beliefs, and in turn, support outgroup censorship:

H3a: The more liberal a participant is, the more they should support censoring conservative speakers, and this should be mediated through their perceptions of the victimization of racial minorities, sexual minorities, religious minorities, and women.

H3b: The more conservative a participant is, the more they should support censoring liberal speakers, and this should be mediated through perceptions of victimization of Whites, Christians, and men.

Method

Sampling

Participants were undergraduates from the Department of Communication subject pool[4]. A starting sample of $N = 284$ was collected through an online survey. Respondents who completed the survey in less than ten minutes ($n = 16$) or responded to too few items ($n = 5$) were removed from the sample. In addition, because PC has some unique forms in the

[4]For all three studies in this dissertation, sign up procedures were organized such that a participant could only be included in one part of the study.

US (and other countries have their own PC ideologies), international students (n = 21) were removed from the sample, leaving a final sample of N = 242. Of these remaining participants, the sample was predominately women (n = 160; 66.1%), and was ethnically diverse: White (n = 96, 39.7%), Asian (n = 51, 21.2%), Latino (n = 47, 19.4%), Black (n = 5, 2.1%), and multiracial or people who do not identify with a listed category; n = 42, 17.4%).

Measures[5]

Perceptions of victimization. In order to evaluate which groups respondents thought were victims of discrimination, they were asked the following: *Please estimate the percentage of each group that you think has been the victim of discrimination based on their group membership during their time at UCSB.* Respondents were asked about members of racial minorities, religious minorities, sexual minorities, and women, as well as about members of majority groups. Specifically, the were asked to estimate the percentage of women, men, Blacks, Latinos, Asians, Native Americans, Whites, Muslims, Jews, Christians, and members of the LGBTQ community that had been victims of discrimination. The results were factor analyzed using Principle Axis factoring with Promax rotation and a loading cutoff of .30. The results indicated a two factor solution. The first factor, *Marginalized victims*, was indicated by seven items that included: women, Blacks, Latinos, Asians, Native Americans, Muslims, and LGBTQ, and formed a reliable scale (M = 40.91, SD = 23.12, α = .94). The second factor, *Dominant victims*, was indicated by three items of groups who are generally thought of as more powerful: Whites, men, and Christians (M =

[5] Items for all three studies can be found in Appendix A

14.63, *SD* = 14.69, α = .64). Jews loaded onto both factors and so the item was not included in either scale.

Support for censorship. Respondents' support for censorship was evaluated by asking them whether or not certain speakers should be allowed to give presentations on campus. Participants read: *Lately there have been discussions at UCSB about which speakers are appropriate to bring to campus. Below is a list of topics that could be discussed at a lecture. Would you support UCSB allowing the following lectures to take place? A speaker who argues that:* The text of each item is listed next to its corresponding factor (see below for items). Topics included both those for which political ideology was expected to be relevant (e.g., climate change, healthcare), and about which ideology was unexpected to be a predictor, which served as distractor items. Participants could answer that they were in favor of allowing such speakers, they were opposed to allowing them, or that they were not sure. These responses were then dichotomized such that a participant not in favor of banning the speaker (includes "yes" and "not sure") (0), or they were in favor of banning the speaker (1). An exploratory factor analysis using principal components analysis, which is better suited for working with dichotomous data, using Varimax rotation and a .30 loading cutoff indicated a four factor solution. One item (*The earth's temperatures naturally ebb and flow, and climate change is not caused by humans*) loaded onto more than one factor and was not included in analyses.

The first factor, *Nutrition*, was made up of two items about food and food labeling (*genetically modified foods are harmful to health*; *the FDA should regulate nutrition labels*). A second factor, *Guns*, only had one item (*The more people who have guns, the lower the crime rate will be*). Because these factors did not represent a particular political ideology (for

the first factor), or only had one item (for the second factor), these factors were not included in further analyses. The third factor, *Ban Liberal Speakers* was made up of three items about topics that would typically be supported by liberals (*the US government should provide health insurance for everyone, the UN should create a global currency*[6]*, there should be a constitutional amendment banning the electoral college and making the president elected by popular vote*[7]; *M* = .10, *SD* = .36). The third factor, *Ban Conservative speakers*, was made up of five items that represented speakers whose positions would typically be supported by conservatives (*Homosexuality is a disease that can be cured through conversion therapy; Transgenderism is a form of mental illness; alt-Right ideology; M* = 1.71, *SD* = 1.15). Composite variables were created out of each factor. Because *Ban liberal speakers* and *Ban conservative speakers* were each indicated by the same number of items, the composite variables were created by taking the sum of the items in the factor, with larger numbers indicating more support for censorship.

Political ideology. Political ideology was measured using two 7-point items that were found to be reliable and predictively valid in other studies (Robinson & Reid, 2015; Robinson & Reid, 2016b), where smaller numbers indicate more liberal views. Participants were more liberal than conservative (*M* = 2.84, *SD* = 1.33). Because the scale was skewed, it was transformed using a logarithmic transformation which improved the distribution. All analyses were conducted with both the original scale and the transformed scale. All

[6] Though this was originally included as a filler item, it loaded onto the *Liberal Speakers* factor. However, at the time of data collection, Donald Trump had made globalization appear to be a liberal issue (Jamrisko, 2016), so it is reasonable that this would factor with other liberal speakers.

[7] During the time of data collection (February-March, 2017), after Donald Trump had won the electoral college vote but not the popular vote, a ban on the electoral college was a common discussion among liberals.

significant effects were the same when using the transformed and original scale, so the original was retained for ease of interpretation.

Results

The hypotheses were tested using General Linear Model (GLM) and Hayes' (2013) Process Macro.

H1: Perceptions of victimhood

The first hypothesis predicted that (H1a) the more liberal a participants was, the more they would perceive victimization among racial minorities, sexual minorities, religious minorities, and women. H1b predicted that the more conservative a participant was, the more victimization they would perceive among Whites, Christians, and men. This hypothesis was tested using a mixed GLM, with political ideology entered as a covariate and *Marginalized victim* and *Dominant victim* entered as a repeated measure. The tests of within-subjects effects indicate that participants perceived that members of marginalized groups were more frequently the victim of discrimination (M = 41.04, SD = 23.08) than members of dominant groups (M = 14.42, SD = 14.17), $F(1, 256) = 471.41$, $p < .001$, $\eta p^2 = .65$. However, the predicted interaction between political ideology and victim group was also significant, $F(1, 256) = 54.87$, $p < .001$, $\eta p^2 = .18$. As can be seen in Figure 1, the more liberal a participant was the more they perceived that members of marginalized groups were victims of discrimination, B = -4.75, t = -4.67, $p < .001$, and the more conservative a participant was, the more they perceived that members of dominant groups were victims of discrimination, B = 2.20, t = 3.32, p = .001. These findings support H1a and H1b.

Follow up test. It also possible that membership in a marginalized group, rather than political ideology, influences perceptions of victimization. In order test for this possibility,

racial group membership (People of Color, n = 163; White, n = 100) and participant gender were added to the GLM[8] for a 2(Racial group membership: People of Color/White) by 2(Participant gender: male/female) by political ideology mixed GLM, with *Dominant victim* and *Marginalized victim* entered as a repeated measure.

The GLM indicated that consistent with results found in the testing H1a and H1b, the interaction between the repeated measure (*Marginalized victim, Dominant victim*) and political ideology remained highly significant, $F(1, 250) = 52.07$, $p < .001$, $\eta_p^2 = .17$. People of Color did not significantly differ from White people on the repeated measure, $F(1, 250) = 1.61$, $p = .21$, $\eta_p^2 = .006$, and participant gender also had no significant interaction with the repeated measure, $F(1, 250) = .007$, $p = .93$, $\eta_p^2 < .001$. All of the higher order interaction effects were nonsignificant (all Fs < 1.64). These findings further support H1a and H1b by indicating that political ideology, rather than membership in a marginalized or dominant group, is predictive of perceptions of victimization.

H2: Support for Censorship

The second hypothesis predicted that (H2a) the more liberal a participant was, the more they would support banning conservative speakers from campus, and (H2b) the more conservative a participant was, the more they would support banning liberal speakers from campus. This hypothesis was tested using a mixed GLM, with political ideology entered as a covariate (mean centered), and *Ban liberal speakers* and *Ban conservative speakers* entered as a repeated measure. The results indicate a main effect across the repeated measure, $F(1, 256) = 540.57$, $p < .001$, $\eta_p^2 = .68$. Participants were more in favor of banning conservative speakers (M = 1.72, SD = 1.21) than liberal speakers (M = .11, SD = .37). The predicted

[8]Participants were not asked about their sexual orientation or religious affiliation. See

interaction between political ideology and the repeated measure was also significant, $F(1, 256) = 29.34$, $p < .001$, $\eta_p^2 = .10$ (see Figure 2). Consistent with H2, the more liberal a participant was the more they favored banning conservative speakers, B = -.21, $t = -4.13$, $p < .001$, and the more conservative a participant was, the more they favored banning liberal speakers, B = .07, $t = 4.29$, $p < .001$.

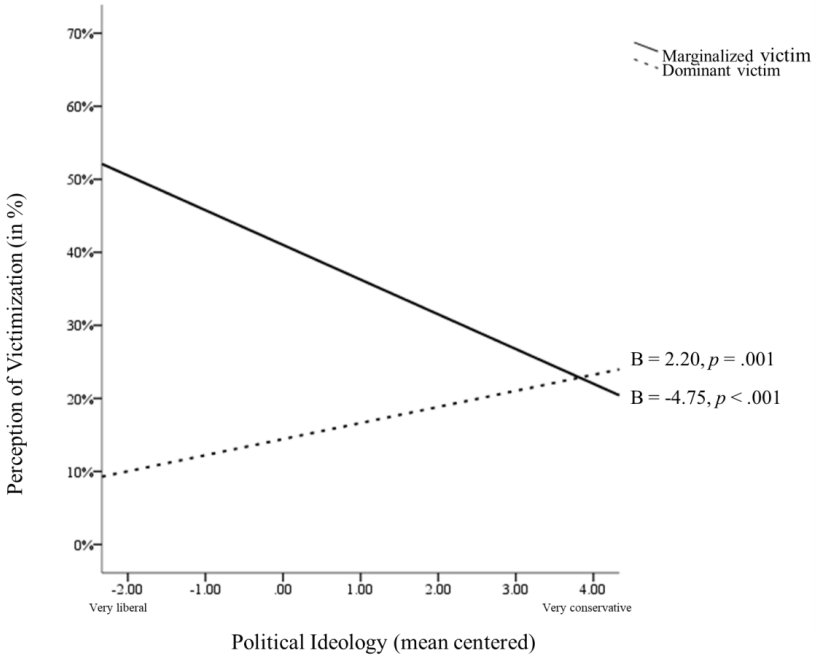

Figure 1. Relationship between political ideology and perceptions of victimization

Follow up test. As in the test of H1, it was possible that membership in a marginalized group could explain support for censorship beyond political ideology. Therefore, a follow up analysis was conducted, with racial group membership and sex added to the model, for a 2 (Racial group membership: White, People of Color) by 2 (Sex: Male, female), by political ideology (mean centered) by repeated measures (*Ban conservative*

limitations in Conclusion (Chapter VI)

speakers, *Ban liberal speakers*) mixed GLM. The results indicate the significant effects from the previous hypothesis test remained. There was a significant main effect across the levels of the repeated measure, $F(1, 250) = 397.02$, $p < .001$, $\eta_p^2 = .61$, and a significant effect of political ideology , $F(1, 250) = 25.04$, $p < .001$, $\eta_p^2 = .09$. The effects of sex, $F(1, 250) = 1.45$, $p < .23$, $\eta_p^2 = .006$, and racial group membership, $F(1, 250) = .60$, $p = .44$, $\eta_p^2 = .002$, were nonsignificant. All higher order interactions were also nonsignificant (all Fs < 2.54), indicating that as predicted in H2a and H2b, political ideology, rather than group membership, is predictive of support for censorship.

H3: Political Ideology, Perceptions of Victimhood, and Support for Censorship

Hypothesis 3a predicted the more liberal a participant is, the more they would support censoring conservative speakers, and this should be mediated through their perceptions of the victimization of racial minorities, sexual minorities, religious minorities, and women (as measured by the composite variable). This hypothesis was tested using Hayes' Process Macro (Hayes, 2013), using model number 4, which is a simple mediational model.

In order to test H3a, political ideology was entered as the independent variable, *Marginalized victim* (mean centered) was entered as the mediator, and *Ban conservative speakers* was entered as the dependent variable. As was found in the testing of H2, there was a significant direct effect of political ideology on participants' desire to censor conservative speakers, such that the lower they were on the political ideology scale (i.e., more liberal), the more they wanted to ban conservative speakers, B = -.25, $t = -4.55$, $p < .001$. The path from political ideology to *Marginalized victim* was also significant, such that the more liberal a participant was (lower on the ideology scale), the more victimization they perceived, B = -

5.21, t = -4.72, p < .001. However, the path from *Marginalized victim* to *Ban conservative speakers* was nonsignificant, B = -.003, t = -1.04, p .30, indicating the perceptions of victimization did not mediate the relationship between liberal political ideology and support for censoring conservative speakers. Therefore, H3a is not supported.

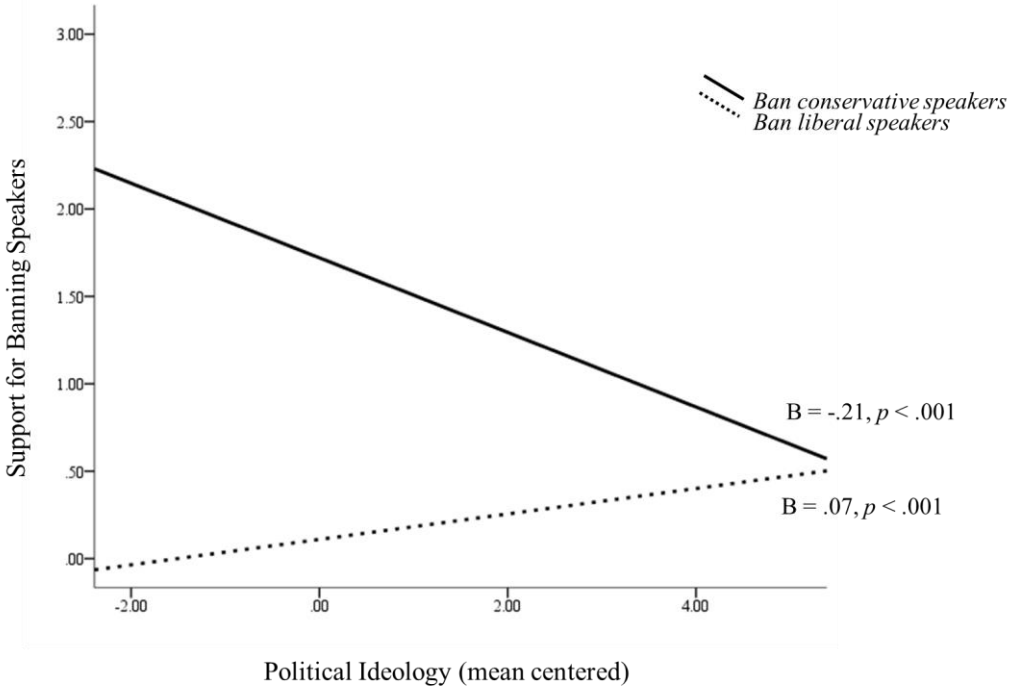

Figure 2. Relationship between political ideology and desire to ban speakers

Hypothesis 3b predicted that the more conservative a participant was, the more they would support censoring liberal speakers, and this would be mediated through perceptions of victimization of Whites, Christians, and men. H3b was tested by entering political ideology entered as the independent variable, *Dominant victim* (mean centered) entered as the proposed mediator, and *Ban liberal speakers* as the dependent variable. The results for H3b were similar to those of H3a. There was a significant direct effect of political ideology on *Ban liberal speakers*, such that the more conservative a participant was, the more they

supported banning liberal speakers, B = .080, $t = 4.59$, $p < .001$. Political ideology was also predictive of *Dominant victim*, B = 2.61, $t = 3.79$, $p < .001$, but *Dominant victim* was not a significant predictor of *Ban liberal speakers*, B = .002, $t = 1.07$, $p = .29$. The findings indicate that perceptions of victimization of Whites, Christians, and men do not mediate the relationship between political conservatism and support for banning liberal speakers. Therefore, H3b is not supported.

Discussion

The purpose of this first study was to examine the extent to which political ideology was predictive of perceptions victimhood and calls for censorship within the context of students' desire to ban potentially offensive speakers from campus. As predicted, there was evidence that liberal and conservative students differed in their perceptions of what groups are discriminated against. More liberal participants perceived that Blacks, Latinos, women, Muslims, members LGBT, and Native Americans were the victims of discrimination, whereas more conservative participants perceived that Whites, Christians, and men were the victims of discrimination. The results also indicate that beliefs about who should be censored can be predicted by political ideology. The more liberal a participant was, the more they supported banning speakers who advocate conservative positions, and the more conservative a participant was, the more they advocated banning speakers who advocate liberal positions. Membership in a marginalized group (based on racial group membership and sex) was not predictive of perceptions of victimhood or support for censorship. Last, contrary to the prediction, the results indicate that perceptions of victimhood do not mediate the relationship between political ideology and support for censoring political rivals, indicating that participants may not have been using their perceptions of victimhood in order to justify their

support of censorship. Taken together, these results indicate that perceptions of victimization and support for censoring political outgroups play an important role in political rivalries.

Political Rivalry and Victimhood

In testing H1, the findings indicate that liberals perceived higher victimization among socially recognized victim groups (minorities and women; i.e., Haidt's seven protected classes (Haidt & Leo, 2017)) than did conservatives, and that more conservative participants perceived greater victimization among groups that have historically held more power in the United States (Whites, Christians, men) than did liberals. These results map onto national surveys that indicate that liberals are more likely perceive victimization among Blacks and conservatives are more likely to perceive victimization among Whites (Jones, 2017).

The results of this study suggest that liberals and conservatives may be engaging in competitive victimhood in the name of their coalitional partners. Research on CV has found that members of adversarial groups have a tendency to claim that they have been victimized by their rivals in order justify actions against them and to avoid paying for past transgressions (Noor et al., 2012). Furthermore, groups may claim victimhood in order to elicit support from outsiders (Young & Sullivan, 2016). The results of this study bring up another possibility: rather than claims of victimhood being used in order to elicit the support from third parties, it also possible that these victim groups are strategically recognized because doing so aids in the building of political coalitions. Liberals have long allied themselves with some, but not all minority groups, and the content of PC norms are specific to these allies (Ford, 2017; Haidt & Leo, 2017; Robinson and Reid, 2016a; see Chapter I). Liberals' selection of which groups deserve victim status is potentially related to which groups are likely to support liberal ideology. For example, liberals advocate for policies to reduce

inequality among Blacks, Latinos, and members of the LGBT community, but expend less energy advocating for such policies for poor rural Whites in Appalachia or the Mississippi Delta, presumably because the former group is more likely than the latter to support politicians and policies on the left (Brownstein, 2017). In the future, research on CV should not just examine how and why groups use claims that their own group has been victimized, but also how groups may use claims of others' victimhood in order to further their own political agenda.

The results of this study also provide important insight into the extent to which perceptions of victimhood are subjective. The members of this sample are remarkably similar in many ways: they all attend the same university and live in the same community. In fact, the item that asked them about their perceptions of which groups are the victims of discrimination were specifically about UCSB. Considering that CV is associated with actions against outgroups and delays in reconciliation processes (Young & Sullivan, 2016), the extent to which victimhood is invoked on the right or left can have important implications for the campus community.

Political Rivalry and Support for Censorship

In testing H2, the results indicate that political ideology is predictive of support for censorship. The more liberal a participant was the more they supported banning speakers who advocated conservative positions, and the more conservative a participant was the more they supported banning conservative speakers. This provides support for the argument made by Robinson and Reid (2016a) that PC is a political tool used to engage in social competition through the control of discourse. The findings presented here also have important practical implications. Many have referenced anecdotal evidence for the fact that decision of who to

censor is often used as a political tool, especially on college campuses (e.g., Al-Gharbi & Haidt, 2017; Delton, 2017; Lukianoff & Haidt, 2015). The results of this study provide empirical support for this argument.

Universities have long been bastions of freedom of speech. A fundamental component of this freedom is that as long as speech is not libelous and does not incite violence, it should be allowed, independent of the political position the speaker advocates (Chait, 2015; D'Souza, 1992; Hentoff, 1992). Participants in this sample, however, appear to have used their political ideology in order to determine what is appropriate for public presentation, and think that their public university should ban speakers accordingly. According to Al-Gharbi and Haidt (2017), the same rules that have allowed conservative speakers to be deplatformed have also been used to ban liberal speakers and fire liberal professors who are accused of espousing liberal views in conservative environments. Though in the short-run, censoring political rivals may be an effective strategy for engaging in intergroup competition, and even temporarily preventing minority groups from being exposed to speech that might offend them, in the long run, allowing decisions of censorship to be determined by political ideology can backfire because it is possible that it inflames intergroup competition, further decreasing the possibility of political reconciliation and progress.

Limitations

There were several limitations in this study. Though there were sampling issues in this study (primarily liberal and female; did not measure sexual orientation or membership in a religious minority), issues of sampling occur in all three studies, so a discussion of this limitation will be postponed until the conclusion in Chapter VI.

Specific to this study is an issue related to the measurement of support for censorship. The prediction made in H3 that participants would justify their support for censorship by using claims of victimhood was not supported. There are possible methodological reasons for why this hypothesis was not supported. First, the items about speakers advocating liberal positions (nationalized healthcare, electoral college, universal currency) were not as controversial, and not as related to identity or victimhood as the items advocated in the conservative items (homosexual conversion therapy, trans-genderism, alt-right ideology). Perhaps if the liberal items had covered topics more related to issues conservatives would associate with victimhood (e.g., affirmative action, redistribution of wealth to underprivileged areas), there would have been greater opposition to those speakers. For those items associated with conservative positions, though they did relate to the victimization of sexual minorities, they did not explicitly relate other marginalized groups (e.g., racial minorities, Muslims, women). The prediction that perceptions of victimization mediate the relationship between political ideology and calls for censoring political rivals relies on groups' abilities to justify censorship with claims of victimization, which means they must be able to claim that those speakers could have been seen as victimizing. Perhaps the mediation predicted in H3 was not supported because it would have been difficult for groups to argue that these speakers were victimizing (other than for sexual minorities). This should be remedied in future replication studies.

Despite these limitations, this study provides insight into how people use their political ideology as a filter through which they perceive victim status and how they determine who should be allowed to speak on a college campus. Robinson and Reid (2016a) argue that PC serves to allow groups to justify censoring outgroups by claiming

discrimination. The results of this study indicate the political ideology is an important factor in determining who is a victim, and who should be censored, meaning people are engaging in identity politics. Therefore, it serves as an important first piece in testing that argument.

Chapter IV: Priming PC Norms and Self-Censorship

Consistent with social identity theory, and research on competitive victimhood, Study 1 showed that concerns regarding censorship are an intergroup phenomenon that involves political identities and beliefs regarding victimhood. Given this direct evidence that PC is a form of identity politics for both the political right and left, an important next step is to examine how PC ideologies are made salient, and what implications this might have for self-censorship. As was discussed in Chapter II, self-categorization theory (Turner, Hogg, Oakes et al., 1987) posits that group members will converge toward their group's prototype as a way of indicating that they are valuable group members, and salient prototypes (e.g., normative attitudes, normative behavior) will shift as different social identities are made salient in different contexts. This study tests the extent to which language can prime different groups' prototypes, which define group norms (see Hogg & Reid, 2006), and is conducted within the context of hypothetical teacher-student interactions in a classroom setting.

As reviewed earlier, there is evidence that conservative code-words (e.g., thug, welfare queen) prime racial group identities, and have effects on political attitudes (Hurwtiz & Peffley, 2005). This chapter tests the analogous idea that common PC terms (such as safe space and microaggression) can also prime political identities and affect political discourse. Unlike conservative code words, which work by making negative stereotypes salient, PC language appears to prime left wing normative attitudes and beliefs (i.e., liberal group prototypes) (Robinson & Reid, 2016b). It is possible that for conservatives (at least on a university campus where PC language and culture is common), PC language acts as a cue that their attitudes are discordant with prevailing social consensus, and they may decide not to risk ostracism by expressing views that are unpopular. Therefore, the goals of Study 2

were three-fold: first, to explore how people take cues from language in order to make inferences about the opinion climate; second, to directly test the hypothesis that conservatives will self-censor in a liberal PC environment; and third, to examine what role individual differences play in the decision to self-censor.

Priming Ideological Support with PC Language Cues

Though PC is an ideology, it also has well known prescriptions about normative language (see Chapter I; Ehrenreich, 1992; Hughes, 2010; Robinson & Reid, 2016a; 2016b). In particular, PC has normative prescriptions regarding the language that should be used to describe some, but not all, minority groups. For example, PC prescribes that people should use person-first language when describing people with disabilities in order to non-essentialize the disability (e.g., person with a disability is PC vs. disabled person; Dunn & Andrews, 2015), and use gender neutral language to remove linguistic sexism (e.g., firefighter vs. fireman; humankind vs. mankind) (for review, see Maass, Suitner, & Merkel, 2013). There are also PC normative prescriptions about group labeling, and these norms often shift over time (Hughes, 2010). For example, it was once appropriate to describe all peoples native to the Americas as Indians. In the 1990s, the preferred term in the United States (i.e., the PC term) became "Native American," which has shifted once again to "American Indian" (Hughes, 2010). Those who support this language claim that group labels change social perceptions of groups, and in turn, decrease discrimination (Collins & Clément, 2012; Sutton, 2010).

Indeed, there is evidence for this argument. Research shows that the language people use to describe groups primes stereotypes, thereby affecting the way listeners perceive the groups being described (Maass et al., 2013). For example, Carnagi and Maass (2007) found

that derogatory labels (e.g., "fag") prime negative stereotypes about homosexuals over and above other labels, such as "gay." However, it is also true that language can be used strategically to prime stereotypes and identity based politics. For example, Hurwitz and Peffley (2005) asked respondents how funding should be spent in order to reduce crime after they had exposed participants to a conservative code-word versus a no code word control (violent *inner city* criminals vs. violent criminals). They found that among participants who endorsed negative stereotypes about Blacks, those in the code word condition were more likely to support spending money on prisons than antipoverty programs, and did so as a function of their endorsement of negative Black stereotypes. Hurwitz and Peffley interpreted their findings as evidence that code words primed negative stereotypes of violent Black men, and this prime was enough for those who endorsed the stereotypes to support policy aimed at reducing this perceived threat.

Just as derogatory labels and conservative code words can be used to prime negative stereotypes, liberal PC ideology has normative prescriptions about language that can similarly prime ideology and stereotypes. For example, though PC code words do not prime negative racial stereotypes, they can still be used to prime race and can be leveraged to encourage support for liberal policies. For example, in a replication of Hurwitz and Peffley's study, Robinson and Reid (2016b) found that when liberals read PC code words (*underprivileged community*) as opposed to conservative ones (*inner city*), they were more likely to perceive that there is unfairness in the justice system, an attitude that is in line with PC ideology. These findings are all consistent with the hypothesis that language choice betrays political ideology (Sutton, 2010), which suggests that for receivers, language choice makes the political positions of speakers salient, and makes political identities salient.

But how does PC normative language affect perceptions of normative attitudes and appropriate discourse? One possibility, alluded to above, is that PC language can affect the perception that there is social consensus regarding liberal political attitudes and beliefs. As would be expected if language is a cue of a speaker's attitudes, hearing a speaker use liberal PC code words (e.g., safe space, trigger warning) is likely to be taken as an indication to listeners that the speaker supports PC ideology, and this should prime political identities and norms. If this is done in a social context where the speaker has authority and social status, as would be the case for a teacher in a classroom setting, this may affect the likelihood that listeners perceive the climate of opinion as more strongly in support of a liberal position than it might otherwise be judged. Indeed, research on the false consensus effect shows that people with relatively extreme opinions are more likely to express them, and this makes it seem like their position has greater social support than it actually does (Miller & Morrison, 2008). If it is true that PC language primes political identities for liberals and conservatives, and if that priming indicates greater consensus around the liberal position, then:

> H1: Compared to non-PC control, PC code words (*diverse set of voice, marginalized, safe space*) should increase perceived normative support for PC protected classes (members of ethnic and sexual minorities) and political liberals but not political conservatives.

Language not only gives listeners an indication of the speaker's attitudes, but it can also give listeners an indication of the norms in that environment (i.e., the opinion climate), and thereby affect self-expression (Collins & Clément, 2012). If PC language primes perceived support for PC norms, then people who identify as liberal are likely to see such cues as evidence that their normative position is dominant, and that they can feel comfortable

expressing their PC-consistent views. For people who do not subscribe to PC, however, PC language may not only act as a prime, but may also amplify the perception that their views are discordant, in which case self-censorship may be a sensible strategy. Therefore:

> H2: There should be an interaction between code word primes and political ideology, such that in the PC code word condition (but not in the non-PC control condition) the more conservative someone is, the more they will self-censor.

Diversity of opinion

The argument of those who claim PC causes self-censorship on college campuses (Chait, 2015; Lukianoff & Hadit, 2015), rests on the assumption that liberal PC is the dominant ideology at most universities, and that by silencing opinions that counter PC normative attitudes (or can be interpreted as opposing them), it necessarily decreases diversity of opinion. However, an alternative argument is often made in support of PC. Namely, PC is not about silencing those who are part of more powerful groups (e.g., White people, men, political conservatives); rather, PC is actually about providing safe spaces for members of marginalized groups to speak (see muted group theory in Chapter I). According to this perspective, by providing these opportunities, PC may actually increase diversity of opinion. Indeed, research has found that members of racial minorities may refrain from participating in class discussions because they feel like they are often asked to speak for their entire racial groups and that they are unfamiliar with the academic discursive style that they feel they must use in the classroom (Rotham, Lispet, & Nevitte, 2003; White, 2011). If this hypothesis is correct, then members of ethnic/racial minorities should interpret PC code words as encouraging them to speak, and they should therefore be more likely to share their opinions:

H3: Members of racial minorities in the PC code word condition should perceive that they are being encouraged to express their opinion more than in the non-PC control condition.

H4: There should be an interaction between code word condition and racial group membership, such that in the PC code word condition (but not in the non-PC control condition) members of ethnic/racial minorities will be more likely to report that they would express their opinion.

Self-Censorship and Individual Differences

The relationships between code words, political ideology, and self-censorship predicted in H2 relies on the assumption that when ideologies are primed, political identities are made salient, and people will self-censor in order to avoid social sanctions. The ways in which fear of social sanctions affects opinion expression has been examined within research on Spiral of Silence (SoS: Noelle-Neumann, 1974). SoS posits that when people perceive that they are in the opinion-minority, they will self-censor, particularly about morally loaded issues (for a more detailed discussion of SoS, see Chapter II). According to SoS, the mechanism that causes self-censorship is the fear of social isolation (FSI) that can result from the social sanctions that one incurs from expressing minority opinions (Noelle-Neumann, 1993).

Not surprisingly, some individuals fear social isolation more than others. These high FSI individuals are more likely to seek information, such as opinion polls, in order to ascertain what the normative opinion is (Hayes, Matthes, & Eveland, 2013) and are more likely to self-censor when they think they are in the minority (Matthes, Hayes, Rojas as al., 2012). A related individual difference, Willingness to Self-Censor (WTSC: Hayes, Glynn, &

Shanahan, 2005) is also predictive of silencing one's opinion when in the perceived minority. The higher an individual is in WTSC, the less likely they are to engage in public political activities such as putting a campaign sign on their lawn (Hayes, Scheufele, & Huge, 2006), and the more likely they are to report that they would avoid sharing their opinion in hostile, but not friendly, opinion climates (Hayes et al., 2005). This is all to say that *if* PC is causing self-censorship on college campuses (which is still a question that needs to be tested), then it likely differentially causes self-censorship for different individuals. PC norms are often unclear (Ford, 2017; Robinson & Reid, 2016a), making it difficult to be certain about the degree of support for minority and majority opinions. Because of this, people who are high in FSI and WTSC should be particularly sensitive to code words as a cue to the opinion climate. Taken together, research on WTSC and FSI indicates that those who are higher on these traits should be more likely to self-censor when they perceive their opinions to be in the minority. If, as was predicted in H1, people use code words to assess the opinion climate, and, as is predicted in H2, self-censor accordingly, then this relationship should be moderated by these individual difference variables. Therefore,

> H5: The relationship predicted in H2 will be moderated by WTSC (H5a) and FSI (H5b), such that the higher a person is in these traits, the more likely they will be to self-censor when they perceive their opinion to be in the minority.

Method

Past research on SoS (e.g., Glynn, Hayes, & Shanahan, 1997) and self-censorship (e.g., Hayes, 2007) has relied on participants' imagining that their opinion is in the minority or majority in a hypothetical situation in order to examine the ways in which people self-censor. This experiment tested the effects of PC using similar hypothetical scenarios, but,

instead of explicitly telling participants the minority/majority status of their opinion, participants were primed with PC language. Participants were randomly assigned to one condition of a 2(Condition: PC code word/control) by continuous moderators (political ideology; willingness to self-censor; fear of social isolation) design.

Participants and Design

Participants (N = 222) were recruited from Communication classes at UCSB and were compensated with a small amount of course credit. Considering that universities tend to be at the forefront of discussions about PC, a college sample was considered appropriate for testing the hypotheses. A small part of the sample was made up of international students (n = 19). Because of the nuanced cultural knowledge of language that is required in order to understand the connotation associated with code words, international students were removed from the sample, leaving a final sample of N = 203. The remainder of sample was primarily women (n = 161; 79.3%) and was ethnically diverse, with the following demographics: White (n = 80; 39.4%), Asian (n = 44; 21.7%), Latino (n = 36; 17.7%), Black (n = 6; 3.0%), or multiracial or did not identify with a listed category (n = 37; 18.2%).

Procedures and Inductions

Participation in this study was completed in two parts. In order to avoid priming participants about the purpose of the study when they responded to the individual difference variables, participants were asked to fill out a questionnaire with the covariates and moderators at least five days prior to their participation in the lab portion of the study. Participants who did not fill out this questionnaire at least five days in advance were not allowed to participate. The two parts of the study were linked with a unique identification number.

Code word induction. Participants were told to imagine that a professor made a statement in class in order to encourage more class participation. They were randomly assigned to read either a statement with several PC code words and phrases (e.g., *I was hoping to hear from a more diverse set of voices*) or to a non-PC control condition (e.g., *I was hoping to have more views expressed*). See the full text of each condition in Table 2.

Measures

Likelihood of giving opinion. In order to measure self-censorship, participants were asked *"How likely would you be to raise your hand to offer your opinion?"* (1 *not at all likely;* 7 *very likely; M* = 3.88, *SD* = 1.75).

***Encouragement of opinion expression*.** Participants were asked which group they thought the professor in the stimulus was encouraging to express their opinion. Participants reported on the extent to which they thought the following groups were being encouraged to speak (1 *not at all;* 7 *very much*)*:* political liberals (*M* = 5.36, *SD* = 1.59), political conservatives (*M* = 4.72, *SD* = 1.82), members of ethnic minorities (*M* = 6.01, *SD* = 1.24), and members of sexual minorities (i.e., LGBTQ) (*M* = 5.77, *SD* = 1.52).

Covariates and moderators.

***Political ideology*.** Political ideology was measured by two 7-point semantic differential items (*My political positions are:* 1 *strongly in favor of Democrats* to 7 *strongly in favor of Republicans*; 1 *extremely liberal;* 7 *extremely conservative*). The items were strongly correlated ($r(192) = .71$, $p < .001$), so the mean of the two items was computed (*M* = 3.06, *SD* = 1.36). The sample was more liberal than conservative.

Table 2.
Text of Code Word Inductions

	Code word induction
PC code word	"I was hoping to hear from a **more diverse set of voices** in this class, and I would especially like to hear from those of you who whose **voices have been marginalized in the past**. Universities must be at the forefront of discussions of social issues, and I hope this classroom can be a **safe space** to have those discussions. I'd like to hear your opinions on the current campus climate."
Non-PC control	"I was hoping to hear **more views expressed** in this class. Universities must be at the forefront of discussions of social issues and it's important to me that everyone knows that they have the right to speak up, **regardless of the popularity of their opinions**. I hope this classroom can be a **place** to have those discussions. I'd like to hear your opinions on the current campus climate."

Note. Text did not appear in bold for participants

Willingness to self-censor. According to Hayes and colleagues (e.g., Hayes et al., 2005; Hayes et al., 2006; Hayes et al., 2010), rather than self-censorship being purely situational (i.e., everyone self-censors when in the minority) as the SoS would suggest, it is instead an individually varying trait. Therefore, WTSC was measured using Hayes et al.'s (2005) 5-point eight item scale (e.g., *There have been many times when I have thought others around me were wrong but I didn't let them know*) (1 *strongly disagree;* 5 *strongly agree*; $M = 3.01$, $SD = .78$, $\alpha = .83$).

Fear of isolation. The SoS rests on the assumption that those who perceive themselves as being in the minority will not share their opinions due to fear of social isolation (FSI). Though this does not measure a communicative process, FSI has frequently been found to be an important moderator of the likelihood of speaking out when in the opinion minority (see Matthes, et al., 2012 for review). Therefore, FSI was measured using Hayes et

al.'s (2011) FSI 5-item FSI scale (e.g., *It is scary to think about not being invited to social gatherings*) (1 *strongly disagree* to 5 *strongly agree*; $M = 3.68$, $SD = .89$, $\alpha = .84$).

Results

H1: Code Words and Perceptions of Acceptable Opinions

The first hypothesis predicted that those in the PC code word condition would perceive that liberals and members of ethnic and sexual minorities were encouraged to speak, and that conservatives were discouraged from speaking. H1 was tested using an ANOVA. The data testing H1 are displayed in Figure 3. The analysis revealed that those in the PC code word condition were more likely to think members of ethnic minorities, $F(1, 202) = 7.32$, $p = .007$, $\eta_p^2 = .04$ ($M_{PC} = 6.24$, $SD_{PC} = 1.15$; $M_{control} = 5.78$, $SD_{control} = 1.28$), were encouraged to express their opinions, and were significantly less likely to perceive that conservatives were being encouraged to express their opinions, $F(1, 202) = 45.42$, $p < .001$, $\eta_p^2 = .19$ ($M_{PC} = 3.95$, $SD_{PC} = 1.74$; $M_{control} = 5.52$, $SD_{control} = 1.55$). The two conditions did not differ in their perceptions that members of sexual minorities were being encouraged to express their opinions, $F(1, 202) = 1.59$, $p = .21$, $\eta_p^2 = .008$ ($M_{PC} = 5.90$, $SD_{PC} = 1.57$; $M_{control} = 5.64$, $SD_{control} = 1.45$). In opposition to H1, those in the PC code word condition were less likely to think that liberals were encouraged to express their opinions, than in the non-PC control condition, $F(1, 202) = 17.45$, $p < .001$, $\eta_p^2 = .08$ ($M_{PC} = 4.92$, $SD_{PC} = 1.73$; $M_{control} = 5.82$, $SD_{control} = 1.28$). Therefore, H1 is partially supported.

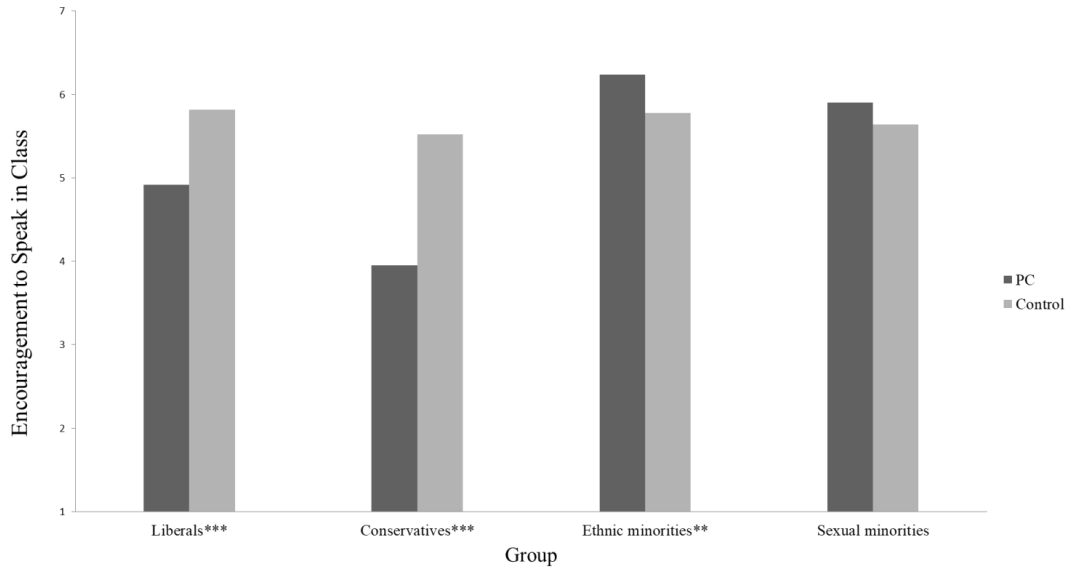

Figure 3. Perceptions of which groups are being encouraged to express their opinions. **Note:** $^{**}p < .01$, $^{***}p < .001$.

H2: Code Words and Self-Censorship

Under H2 it was predicted that there would be an interaction between political ideology and code word condition, such that more conservative participants (but not liberals) would be more likely to report that they would self-censor in the PC code word condition than in the control condition. H2 was tested using a 2(Condition: PC code word/control) by political ideology GLM, predicting likelihood of offering one's opinion. The GLM showed that the main effects of condition, $F(1, 188) = 1.38$, $p = .24$, $\eta_p^2 = .007$, and political ideology, $F(1, 188) = .47$, $p = .49$, $\eta_p^2 = .003$, were nonsignificant. However, the predicted interaction between condition and political ideology was significant, $F(1, 188) = 6.02$, $p = .015$, $\eta_p^2 = .031$ (see Figure 4). Follow up simple slopes analyses indicated that political ideology (mean centered) was not a significant predictor of willingness to offer one's opinion in the control condition ($\beta = .13$, $t = 1.23$ $p = .22$), but was a significant predictor of likelihood of sharing one's opinion in the PC code word condition, such that the more

conservative someone was, the less willing they were to share their opinion (β= -.23, t = -2.26 p = .025). Therefore, H2 is supported.

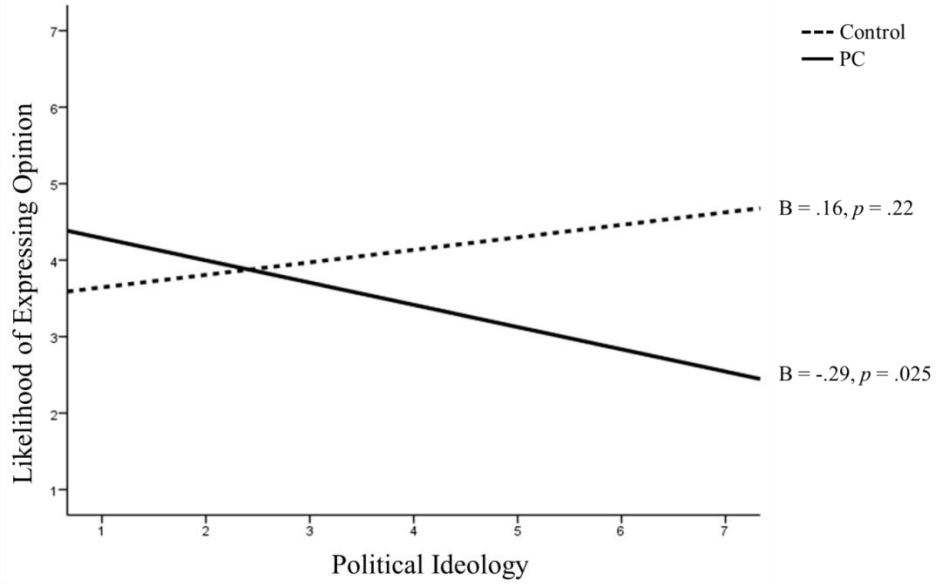

Figure 4. Interaction between code word condition and political ideology on likelihood of expressing opinion

H3 & H4: Marginalized Groups and Opinion Expression

H3 predicted that People of Color in the PC code word condition would perceive that they were being encouraged to speak up more than People of Color in the non-PC control condition. In order to test this hypothesis, participants were categorized as either White (n = 80), or People of Color (n = 123). This hypothesis was tested using a 2(Condition: PC code word/control) by 2(Racial group membership: People of Color/White people) GLM predicting perceptions that members of ethnic minorities are being encouraged to speak. As in the test of H1, the results of the analysis reveal that condition was a significant predictor of perceptions that members of ethnic minorities were being encouraged to speak, $F(1, 199) = 5.75$, $p = .017$, $\eta_p^2 = .028$. The effect of racial group membership was nonsignificant, $F(1, 199) = 1.77$, $p = .19$, $\eta_p^2 = .009$. The interaction between condition and racial group was

nonsignificant, $F(1, 199) = 2.56$, $p = .11$, $\eta_p^2 = .013$, indicating that White people and People of Color both interpreted the message as encouraging members of ethnic minorities. Therefore, H3 is not supported.

H4 predicted an interaction between condition and ethnic group membership on likelihood of expressing one's opinion, such that People of Color would be more likely to report that they would express their opinion in the PC code word condition than in the control condition, controlling for political ideology. H4 was tested using a 2(Condition: PC code word/control) by 2 (Racial group membership: People of Color/White people) GLM with political ideology entered as a covariate, predicting likelihood of expressing one's opinion. There was no main effect for political ideology, $F(1, 187) = .092$, $p = .76$, $\eta_p^2 < .001$. There were marginally significant main effects for condition, $F(1, 187) = 3.32$, $p = .070$, $\eta_p^2 = .017$, and ethnic group membership, $F(1, 187) = 2.84$, $p = .093$, $\eta_p^2 = .015$. The predicted interaction between condition and ethnic group membership was significant, $F(1, 187) = 7.44$, $p = .007$, η_p^2 .038. However, the direction of the effect was inconsistent with H4. Though People of Color ($M = 3.89$, $SD = 1.61$) were slightly less likely than White people ($M = 4.14$, $SD = 1.75$) to report that they would express their opinions in the control condition, the biggest difference was between the likelihood of the two groups expressing their opinion in the PC code word condition (People of Color: $M = 4.13$, $SD = 1.72$; White: $M = 2.97$, $SD = 1.83$) (See Figure 5), meaning that People of Color were equally likely to speak up in both conditions, but White people were discouraged from speaking up in the PC condition. Therefore, H4 is not supported.

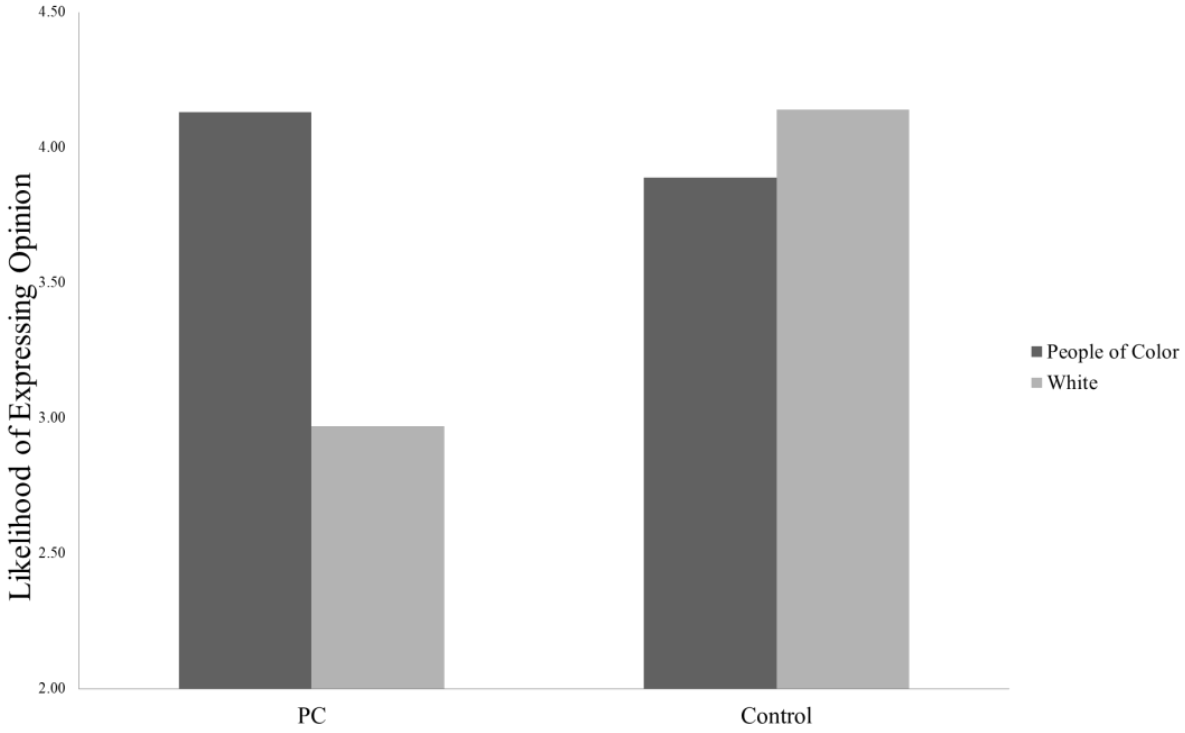

Figure 5. Interaction between condition and ethnic group membership of the participant predicting likelihood of sharing their opinion

H5: Individual Differences

H5 predicted that the condition by political ideology interaction found in H2 would be moderated by (H5a) WTSC and (H5b) FSI. H5a was tested using a 2 (Condition: PC code word/ control) by political ideology (mean centered) by WTSC (mean centered) GLM predicting likelihood of expressing one's opinion. The results of the GLM show a significant main effect of WTSC, $F(1, 184) = 11.73$, $p = .001$, $\eta_p^2 = .060$, such that the higher a participant was in WTSC, the less likely they were to report that they would offer their opinion ($\beta = -.24$, $t = -3.44$, $p = .001$). The interaction between condition and political ideology found in H2 remained significant, $F(1, 184) = 5.10$, $p = .025$, $\eta_p^2 = .027$. All other effects were nonsignificant (all $F < .72$), including the predicted three-way interaction between condition, political ideology, and WTSC, $F(1, 184) = .04$, $p = .84$, $\eta_p^2 < .001$. Therefore, H5a is not supported.

H5b was tested using a 2 (Condition: PC code word/control) by political ideology (mean centered) by FSI (mean centered) GLM. The results of the GLM showed a significant main effect of FSI, $F(1, 183) = 4.22$, $p = .041$, $\eta_p^2 = .023$, such that the higher a participant was in FSI, the less they likely they were to report that they would share their opinion ($\beta = -.37$, $t = -2.06$, $p = .04$). The interaction between condition and political ideology was also significant, $F(1, 184) = 4.46$, $p = .036$, $\eta_p^2 = .024$. There were no other significant effects (all $F < 1.42$), including the predicted three-way interaction between condition, political ideology, and FSI, $F(1, 183) = .06$, $p = .81$, $\eta_p^2 < .001$. Therefore, H5b is not supported.

Discussion

Study 2 provided insight into the decades-long debate about the extent to which PC encourages self-censorship. First, the test of H1 showed that PC language primed the perception that liberals, conservatives, but not People of Color, were being discouraged from expressing their opinions. Second, the test of H2 indicated that in the PC code word condition, the more conservative a participant was, the more likely they were to report that they would self-censor. Third, tests of H3 and H4 found that condition, rather than racial group membership, was the best predictor of how participants interpreted the message: both White people and People of Color interpreted the message as encouraging minority group. Subsequently, when asked if they would share their opinion, People of Color in the PC condition were not more likely to report that they would speak up than those in the control condition, but White students exposed to the PC condition were less likely to report that they would speak up than those in the control condition. And fourth, the test of H5 found only main effects for the individual differences of WTSC and FSI, indicating that for those high in these traits, they may self-censor independent of the opinion climate.

Language and Ideological Primes

The findings of this study provide some support for the argument put forth by those who have argued for the harmful effects of PC (Berman, 1992; Chait, 2015; Hentoff, 1992; Lukianoff & Haidt, 2015; Robinson & Reid, 2016a; Schlosser, 2015): if using PC language was enough to prime the ideology, then this may indicate that PC ideologies have authority on campus, and may indicate to students that expressing a PC inconsistent view or inadvertently violating a PC norm is not acceptable. The findings of this study indicate that priming PC using ideological normative language may be enough to as a cue of what a normative attitude is, and that views inconsistent with PC should not be expressed. It has been well documented that subscribers of different political ideologies use different words to frame the same issue (e.g., *illegal* vs. *undocumented immigrants*; Demby, 2013). This study suggests that those who use these kinds of terms are validating their ideological positions, and that listeners infer those positions, and react accordingly. Furthermore, in line with the original intent of those in support of PC, these findings show that PC language is interpreted as encouraging (i.e., creating a friendly environment) for members of ethnic minorities to speak up, but does not necessarily result in them speaking up above and beyond the use of non-PC language. Furthermore, the test of H4 indicates that PC code words can be interpreted as creating a hostile opinion climate (i.e., indicates that they are in the opinion-minority) for White students, including those who subscribe to a liberal ideology. Though on its face this may seem positive, as encouraging members of ethnic minorities to share their views can benefit everyone in the classroom, there are tradeoffs to everything, and in this case, participants also interpreted PC language as discouraging for political conservatives. It

may be that these PC primes prevent students from hearing variety in political positions in a classroom setting.

Past self-censorship research has found that people self-censor when they feel their opinion is in a minority. Study 2 shows that language alone can prime ideology, indicating to non-adherents that they are in the opinion minority and should self-censor. Political conservatives are the minority both in the sample and on the campus where the data were collected. The findings indicate that the more conservative a participants was, the more likely they were to self-censor in the PC condition. Due to the small number of truly conservative students in the sample, many of those students who reported that they would self-censor are probably political moderates, or even (arguably) political liberals, but are just less liberal than the rest of the sample. This is consistent with the argument made by Loury (1994), who claims that moderates are also likely to self-censor in a PC environment so as to not risk looking like a deviant community member. Though more research is needed, the findings here indicate that PC has a certain amount of ideological authority on campus, such that priming that ideology through the use of language, even for moderate liberals, was seen as a cue that dissenting views should not be expressed.

Minority Groups and Opinion Expression

One of the main arguments of those in favor of PC is that its prescriptions are necessary in order to encourage those who come from groups that have been historically been the discriminated against. According to this line of thinking, even if some voices from dominant groups are silenced, diversity of opinion should rise overall because more people from more diverse backgrounds will be speaking up (e.g., Orbe, 1998; White, 2011). The data from this study, however, indicate that People of Color were equally likely to report that

they would share their opinion in the PC condition as the non-PC control, meaning the PC prime was, perhaps, unnecessary in order to encourage them to speak. Whites, on the other hand, were less likely to report that they would share their opinions in the PC code-word condition as the non-PC control condition. Taken together, these findings indicate that the PC language prime did not encourage members of marginalized groups above and beyond the control condition, but did discourage White students.

There are several ways to interpret these findings. On the one hand, they show that it is possible for instructors to frame sensitive topics in a way that encourage people from different groups and ideologies to express their opinions. However, because the study design asked participants how they would respond to the instructor, and not how they would respond given different reactions from classmates, these data do not make it possible to make strong claims about the effects of PC on diversity of opinion. Importantly, these data are not prescriptive: it is up to each instructor to decide what is best for their classroom. But, at a minimum, instructors should be aware of the effect these code-words can have on the classroom environment.

Self-Censorship

This study has also made contributions to the study of self-censorship broadly, and the SoS specifically. Past experimental research on self-censorship has largely manipulated the opinion climate by asking participants to imagine that their own opinion is in the minority or majority (c.f. Hayes et al., 2010), or by showing them opinion polls, thereby explicitly indicating what the normative attitude is (Hayes et al., 2011). However, in reality, people are rarely explicitly told what majority and minority opinions are, and instead rely on social and environmental cues in order assess the opinion climate.

A fundamental assumption of SoS is that people have a quasi-statistical sense, and that they use this sense in order to assess the opinion climate. However, because past research has relied on these more explicit operationalizations of the opinion climate, this notion of the quasi-statistical sense has gone untested. This is important because assessing social and environmental cues can lead to more uncertainty about the opinion climate than do these explicit operationalizations. If, as Noelle-Neumann (1974) claims, people self-censor when in the opinion minority in order to avoid the social isolation associated with being viewed as deviant, then in reality, without explicit cues about minority/majority opinions, people may be likely to err on the side of caution. Essentially, they may be more likely to interpret social cues as a sign that the opinion climate is discordant with their own views. The results of the current study indicate that participants use social cues in order to assess the opinion climate. In particular, language can prime ideology, and thus what views are and are not acceptable within that environment. Future tests of SoS should examine the quasi-statistical sense in order to assess the effects of uncertainty about the opinion climate.

Individual Differences

This study also contributes to our understanding of individual differences in willingness to self-censor and the fear of social isolation. Research on WTCS (Hayes, 2007; Hayes et al., 2005; Hayes et al., 2006) and FSI (Hayes et al., 2013; Matthes et al., 2012) has found that both individual differences should be predictive of opinion suppression in hostile – but not friendly – opinion climates. The findings presented here indicate only main effects for WTSC and FSI, and there was no evidence for interactions between political ideology and code word conditions. Essentially, people high in WTSC and FSI were likely to self-censor regardless of the minority/majority status of their opinion. It may be that PC topics are so

sensitive that those who are already predisposed to self-censor will do so unless they have absolute certainty that they are in a friendly environment. When only using language to prime ideology, rather than giving participants explicit information about the opinion climate, they were not able to have this absolute certainty.

In fact, when Hayes (2007) told participants to imagine that they were either in the minority or majority opinion at a party, and they were directly asked their opinion on one of three topics (death penalty, affirmative action, or environmental extremism), participants' willingness to express their opinions differed in the friendly and hostile conditions when the topic was death penalty and environmental extremism, but did not differ in the affirmative action condition (a PC-salient topic). Hayes did not attempt to explain this result, however one possibility is that participants were uncomfortable enough with the topic of race that the friendliness of the environment was irrelevant. For those who have a high fear of social exclusion, it may be safer to assume that a classroom is a hostile—rather than friendly—environment, meaning it is safer to self-censor rather than risk social sanctions for saying something inconsistent with PC ideology. It may be that FSI and WTSC are more likely to have main effects (rather than interact with opinion climate) when it comes to topics about which it is easy for a person to inadvertently violate normative attitudes and behavior.

Limitations

Despite the important implications of the findings here, there are some limitations of this study. First, participants may not have perceived the code words in the PC condition as being PC. As was discussed in Chapter I, people have many different definitions of PC, and along with that, differing interpretations of what qualifies as language or an opinion that it consistent with the prescriptions of PC. Though the PC code word manipulation was

perceived as encouraging groups that are generally seen as protected classes by contemporary PC (racial minorities), there was no explicit manipulation check that examined the extent to which participants viewed the language as PC. Follow-up studies should include these manipulation checks. An expanded discussion of this limitation is included in Chapter VI.

Second, the participants in this study were largely liberal, as is the university where the data were collected. It may be that within this particular context, participants already perceive that PC consistent attitudes are normative, so PC primes only reiterate moderates' and conservatives' perceptions of being in the opinion-minority. It is plausible that in an environment in which conservatives make up a larger percentage of the population, PC language might still indicate that the speaker endorses PC ideology, but it may prime intergroup competition and an opportunity to challenge the speaker, rather than produce self-censorship. Future research should examine the effects of ideological primes within different political climates. Despite this limitation, considering the growing concerns over the extent to which liberal PC is dictates policy and decision making on many college campuses (e.g., Kipnis, 2015; Lukianoff & Haidt, 2015; Schlosser, 2015), the tests of these questions within a liberal environment still provide important insight.

Third, like many past studies about self-censorship, this study presented participants with a hypothetical scenario. As Hayes et al. (2005) point out, though we can still gain important insights from what participants report that they would do, more can be gained from studying their actual behavior. Therefore, in an attempt to remedy the limitations to the conclusions that can be drawn from self-reported hypothetical responses, Study 3 will test the questions posed here through behavioral measures.

Chapter V: PC Primes and Political Deliberation

The findings of Study 2 showed that language primes ideology, and people attend to these cues in deciding whether or not to express their opinions. Specifically, more conservative participants and White people reported being more likely to self-censor after being primed with PC code words compared with a non-PC control, whereas People of Color maintained an equal likelihood of self-expression under a non-PC control and PC code word conditions. These findings indicate that PC may have the effect of discouraging expression among Whites and more conservative participants, without a corresponding increase in the opinion expression of minorities. These findings are of practical importance because self-censorship likely decreases public deliberation, and in the classroom setting, it can prevent students from being exposed to a variety of viewpoints and learning how to constructively engage with those who disagree with them. But more than this, as argued in Chapters I and II, it is likely that PC does not merely affect rates of expression, but it may also affect the quality of expression.

Study 3 has five aims and will be conducted within the context of a dyadic conversation between a participant and a research confederate about the topic of banning potentially offensive speakers. The first aim is to replicate Study 2 using social interaction rather than hypothetical scenarios. The second purpose is to test a specific form of self-censorship; specifically, whether conservative students conform to liberal positions under conditions when they are exposed to a PC confederate. It is possible that apparent conformity is an expedient strategy for conservatives to avoid rebuke and sanctions when they find themselves in a PC social situation. The third aim was to analyze the content of participants' communication when they interacted with confederates under PC and non-PC conditions. To

do so involves analyzing large quantities of text, and this can be done efficiently using automated, computer techniques. The fourth aim was to test whether the quality of argumentation under PC and non-PC conditions differs. Specifically, the quality of argumentation can be assessed by measuring the extent to which integrative complexity changes with political ideology and the salience of PC norms. Relatively low integrative complexity involves low differentiation among issue positions, and a failure to reconcile different positions. Given that PC involves polarized identities, it is possible that integrative complexity is lessened among political liberals who invoke PC when discussing political issues, particularly under conditions where PC norms are already salient. Finally, this study tested whether individual differences in willingness to self-censor and fear of social isolation have any independent or interactive effects with PC versus non-PC conditions, and differences in political ideology. In what follows, hypotheses regarding self-censorship, integrative complexity, and individual differences are described.

Conformity, PC and Self-Censorship

If PC norms do lead conservatives to self-censor, as was indicated in Study 2, then it should be possible to observe self-censorship behaviorally in interpersonal discussions. There are several ways in which self-censorship might emerge in interpersonal discussions. People with non-PC attitudes may self-censor by conforming to the majority position (Asch, 1956), stay silent (e.g., see Glynn, Hayes, & Shanahan, 1997 for review), or they may avoid expression by asking counter questions, expressing ambivalence, or referring to someone else's opinion (see Hayes, 2007). In the current study, political discussion will occur in dyads, so a likely avoidance strategy would be to conform to a PC speaker's views. If it is true that PC affects opinion expression by leading people to conceal their opinions for fear of

being judged politically incorrect, then we should find that being presented with a PC consistent position by a confederate leads conservatives to conform to a PC position more readily than in a situation where the confederate presents a non-PC position.

H1: The more conservative a participant is, the more likely they will be to self-censor by conforming to their partner's PC position than a non-PC partner, but only when exposed to a PC normative position.

Language Content and PC

Rather than simply asking people how likely they would be to express an opinion (as was done in Study 2 and is common in self-censorship research), it is also possible that self-censorship may be betrayed by the content of the language that people use. Natural language has frequently been used to draw inferences about communicators. The analysis of natural language has been used in clinical settings to diagnose psychological disorders (Pennebaker & Graybeal, 2001) and by social psychologists and communication scholars to evaluate personality (Chung & Pennebaker, 2008; Gallois & Pittman, 1995; Potter & Wetherell, 1995). More recently, with the rise of computer aided textual analysis (CATA), analyzing texts became markedly less labor intensive (Ireland & Mehl, 2014), and there are now hundreds of published studies using CATA methods to assess personality and other traits in unobtrusive ways that are consistent with standard self-report measures. Research has shown that these techniques work across levels of analysis. For example, individuals exhibit language idiosyncrasies (i.e., "linguistic fingerprints") (Pennebaker & Graybeal, 2001), personality traits are reflected in linguistic styles (for a review of natural language indicators of personality traits, see Ireland & Mehl, 2014), and group processes can be mapped through linguistic style matching where communicators match each other's use of style words (e.g.,

pronouns, prepositions)(e.g., Danescu-Niculescu-Mizil, Lee, Pang et al., 2012; Ireland & Penebaker, 2010). Though the automated coding of texts is not a panacea, as computers do not *interpret* language (see note in Discussion section if this chapter), the automated analysis of natural language is proving to be a valid tool across a range of attributes and levels of analysis. What the technique loses in nuance, it makes up for with sample size and convenience, and it remains predictively valid.

CATA techniques are well suited to studying political discourse, and have been used extensively in political science in order to examine partisan differences in language and themes in political speeches (e.g., Grimmer, & Steward, 2013; Laver, Benoit, & Garry, 2003; Monroe, Colaresi, & Quinn, 2008). Though it may be possible to come up with *some* of the language that participants might use in a conversation about a PC salient topic, these pre-defined word lists often miss important idiosyncrasies in natural language (Chung & Pennebaker, 2008). For example, when a researcher places a word on a pre-defined list, a participant can use a synonym of that word, and it would be missed by dictionary based automated language analyses. Furthermore, a conversation about a PC-relevant topic--in this case, whether or not to ban speakers from campus if they might offend minority groups--can take many different directions and cover many topics. As such, a more flexible approach is preferred in this context.

The Meaning Extraction Method (MEM) (Chung & Pennebaker, 2008) is an inductive form of topic modeling that extracts word categories, from which broader themes can be inferred. The MEM assumes that words frequently co-occur into meaningful categories, and researchers can inductively interpret the meaning of these categories (Boyd, 2017a; 2017b; Chung & Pennebaker, 2008). After inputting a text corpus, the researcher is

left with a binary dataset indicating whether a participant used a word or not (usually just the most frequent content words). These binary data can then be analyzed using a Principal Components Analysis (PCA) in order to extract word clusters that are specific to the content of the corpus of interest (unlike dictionary-based approaches, which contain relatively fixed dictionary word lists). The researcher can then evaluate which texts in the sample contain each factor/theme and can get an understanding of which writers (or experimental conditions) are invoking which themes, how participants are thinking about themselves in relation to the topic of interest (Chung & Pennebaker, 2008), and can then include the use of these themes as variables in subsequent analyses (Boyd & Pennebaker, 2015; LeFebvre, LeFebvre, Blackbrun, & Boyd, 2015). Using the MEM, researchers have found that naturally occurring patterns in language are correlated with personality traits (Pennebaker & Chung, 2008), personal values (Boyd, Wilson, Pennebaker et al., 2015), and students' assessment of their own school performance (LeFebvre, et al., 2015).

If it is true that the language people use to discuss topics can provide meaningful insight into how participants think about these topics and what they find important, then analyzing conversations about a PC-relevant topic should provide insights not only into how participants use PC normative language, but also how they discuss PC topics. Therefore, I propose the following research question to be investigated empirically using the MEM technique:

RQ1: What communication patterns emerge during conversations about PC topics?

Integrative Complexity

According to deliberative democracy scholars, true deliberation requires that people not only express their ideas rationally, but also that they be open to the possibility of

changing their mind when presented with new information (Jackman & Sniderman, 2006). This requires that deliberators consider alternative positions and weigh them against their own perspective. One way to examine the extent to which an argument is embodying the ideals of deliberative democracy is the extent to which it exhibits integrative complexity (IC). IC is an operationalization of cognitive complexity, a cognitive style that is characterized by a higher tolerance for uncertainty and an ability to weigh complex tradeoffs against each other (Tetlock, 1983). Thinking/arguments that exhibit higher IC are represented by the differentiation of viewpoints and their integration into an argument (e.g., Streufert, Suefeld, & Driver, 1965; Suedfeld, Tetlock & Steufert, 1992). Differentiation refers to the extent to which a person not only recognizes that different viewpoints exist, but is also able to understand the subtleties of how viewpoints are different from each other as well as what similarities exist. Integration refers to the extent to which a person is able to weigh the costs and benefits of different options against each other, possibly to offer a solution that transcends or resolves conflicts between the opposing viewpoints (though people rarely reach this highest level of IC).

Though many studies have found that liberals tend to make arguments that have higher IC than conservatives (e.g., Brundidge, Reid, Choi, & Muddiman, 2014; Golec de Zavala, Cisak, & Wesolowska, 2010; Tetlock, 1983; 1985), this is not always the case. For example, Tetlock (1986) examined IC in the context of value pluralism (e.g., the issue of CIA surveillance of personal communications pits two values against each other: personal freedoms vs. national security), and found that people exhibit higher IC when they are trying to reconcile conflicting values compared to those who definitively prioritize one value over another. Research has also found that there are contexts in which political liberals exhibit

lower IC than conservatives. For example, Conway, Gornick, Houck et al. (2016) found that liberals made less integratively complex arguments than did conservatives on topics such as the death penalty, access to birth control, and immigration policy. Conway et al. attribute this finding to the fact that liberals, like conservatives, have certain political issues about which they rely on dogma or ideology, meaning they see these issues in simple good/bad terms, rather than considering the complexity of the issue and the possibility that the other side has a valid point of view.

According to Robinson and Reid (2016a), one of the features of PC is that a failure to support the PC position is attributed to a moral failing rather than a reasonable difference of opinion. This observation suggests that PC ideologues are unlikely to recognize the legitimacy of alternative political viewpoints; it is possible that those viewpoints will simply be ignored or discounted. If this is the case, then PC relevant topics likely fall into the domain of issues about which more liberal participants exhibit lower IC than more conservative participants. Furthermore, because the PC condition frames the issue of banning speakers in terms of values that are relevant to PC ideology (specifically protecting minority groups), but the non-PC condition frames the issue in terms of freedom of speech[9], which should not prime PC ideology, there should be an interaction between condition and political ideology such that liberals express lower IC in the PC condition. Because freedom of speech is not a domain about which conservatives are ideologically committed (Conway et al., 2016), the IC of more conservative participants should not be affect by condition:

[9] Though freedom of speech has recently become an issue that has been associated with White supremacy groups, and thus now may be politically salient for those who support PC ideology, this was not the case at the time of data collection.

H3: The more liberal a participant, the lower their IC will be in the PC condition than in the non-PC condition. The IC of conservative participants should not differ across conditions.

Individual Differences

Study 2 found that people who were higher in willingness to self-censor (WTSC) and fear of social isolation (FSI) were more likely to report that they would self-censor, but these effects were not moderated by political ideology or the PC code-word induction. These main effects of WTSC and FSI fail to confirm past research (Hayes et al., 2005; Hayes, Matthes, & Eveland, 2011), which has emphasized that WTSC and FSI are most predictive of self-censorship when a person is in the opinion minority than the majority (e.g., Hayes et al., 2005). Given that Study 2 involved self-reported self-censorship, it is possible that any effects of climate of opinion and group membership variables were not sufficiently powerful to induce an effect. Study 3 addresses this limitation of Study 2 by testing the hypothesis that WTSC and FSI are responsive to primes in a situation where participants engage in an actual discussion.

H4: In the PC condition, there will be interaction between political ideology and WTSC, such that for conservative participants, the higher a participant is in WTSC, the less engaged they should report having been in the conversation.

H5: In the PC condition, there will be interaction between political ideology and WTSC, such that for conservative participants, the higher a participant is in FSI, the less engaged they should report having been in the conversation.

Method

The hypotheses and research question were tested in a 2(Condition: PC/non-PC) by political ideology by individual difference variable (i.e. WTSC and FSI; for H4 and H5 only) experiment. Because language use was of primary interest, the participants engaged in a text-based conversation with a confederate. This allowed for the analysis of the conversation without the need for transcription.

Participants and design. Participants (N = 168) were undergraduate students enrolled in lower division Communication courses at UCSB. Because of the focus on language as a dependent measure, international students, who vary in their English abilities, were removed from the sample (n = 13), leaving a final sample of N = 155. The sample was predominantly female (n = 135; 87.1%), and was racially diverse, with following makeup: White (n = 58; 37.4%), Asian (n = 38; 24.5%), Latino (n = 29; 18.7%), Black (n = 5; 3.2%), and multiracial or did not identify with a listed category (n = 25; 16.1%). Participants were assigned to either a PC condition (in which a confederate supported banning speakers who may offend minority groups) or non-PC condition (in which a confederate opposed banning speakers even if they offend minority groups). Potential moderators were measured (i.e., political ideology, WTSC, and FSI) during a pre-measures survey that participants took before coming to the lab. Participants completed the in-lab portion of the study in private rooms (see procedures below). Because of this, two people participated at one time, and a research assistant engaged in an online chat with both of them simultaneously. To simplify procedures and allow the research assistant to be able to maintain the same role with both participants (PC or non-PC), participants were assigned to conditions based on time of day, rather than by true random assignment. Those who participated in an even hour of day (e.g., 12:00, 2:00) were assigned to the non-PC condition (n = 79) (for more information on

conditions, see Inductions) and those who arrived at an odd hour (e.g., 9:00, 1:00) were assigned to the PC condition (n = 76).

Procedures. Participation took place in two parts. After signing up for the study, participants were sent a link to an online survey where they completed individual difference measures and demographic items. Five or more days later, participants came to the lab to participate in the second part of the study. Upon arrival to the lab, a research assistant ushered participants to a room where they completed the study in a private cubicle. Participants read a consent form that told them that the researchers were interested in how people discuss important issues differently across various modalities. The research assistant told each participant that another department on campus was also participating in this research, and that they would discuss a topic related to university life with another participant in the other department. In reality, participants had a computer mediated conversation with a confederate (the research assistant), and all participants discussed the same topic: banning speakers who may offend minority groups from campus. Participants were also told that they would be randomly assigned to have a text based chat or a video call. This was done in order to make participants think that they might actually have to speak with and be seen by their partner at some point. Participants' communication with confederates was done using Gmail and Google Hangouts accounts that were created for the study.

After participants read the instructions (see below), the research assistants told participants that they would receive an email from their communication partner and that they should read the email and then take 10 minutes to respond. After 10 minutes, the research assistant returned and told participants to send their email back to the communication partner. Participants were then told to click forward in their survey in order to find out if they were

assigned to a video or text based chat. The research assistant then opened Google Hangouts. The research assistant then closed the door and returned to their own computer where they had a 15 minute chat with the research participant. After 15 minutes, the research assistant returned to the participant's room and told them to end the chat. These chats were not included in the automated language analyses presented in this dissertation, but will be manually coded in future research. Participants then filled out a series of measures about their opinions of the chat and their perceptions of their communication partner. At the conclusion of the experiment participants read a debriefing that explained the full purpose of the study. The research assistants were instructed to make sure that participants understood that they were talking to a confederate and not a real participant before they left, and many participants verbalized how surprised they were to find out that it was not a real participant.

Inductions.

Prompt. All participants discussed the same topic and read the following prompt before reading the email sent to them by the confederate:

> *Lately there has been a lot of talk about controversial speakers coming to college campuses. Some people think that speakers who are offensive towards marginalized groups should not be allowed to speak on campus, whereas others think that universities need to protect the freedom of expression over all else. What are your thoughts on this topic?*

This topic was chosen because of its relevance at the time of data collection. Data collection began one week after violent protests broke out at UC Berkeley in response to the scheduled lecture by Milo Yiannopoulos, so it was assumed to be a salient issue.

PC prime. In order to establish the normative attitude, the confederate sent the initial email message. Participants read one of two emails: either a PC consistent opinion or a non-PC opinion. Based on the suggestion of the research assistants (who are students at the same university as the participants), the email was written in full sentences with proper grammar. The research assistants suggested that this type of language would be expected in an email, and that more informal language should be used in the chat portion. In the PC condition, the email was interspersed with PC code words (*safe space, racist, misogynistic, homophobic, marginalized groups*) and advocated the position that allowing speakers on campus who offend minority groups would ultimately decrease diversity of opinion. In the non-PC condition, the email did not use PC code words, and advocated the position that banning speakers who offend minority groups would ultimately decrease diversity of opinion. See Table 3 for the full text of both emails.

Table 3.
Experimental Induction (Email)

Condition	Text
Pro ban	College campuses should be a **safe space** where everyone can feel comfortable living and learning. When the university allows **speakers who preach racist, misogynistic, or homophobic views**, it creates a chilling effect where **members of marginalized groups** will no longer feel comfortable expressing their opinions. Diversity of opinion in the campus community benefits everyone, and the best way to ensure that diversity exists is to make sure people from all backgrounds feel comfortable speaking up. Essentially, **by allowing speakers tocome to campus who make marginalized groups feel threatened**, the university would be allowing the freedom of speech of one group to override the freedom of speech of another.
Oppose ban	College campuses should be a **place** where everyone can feel comfortable living and learning. When the university **allows others' offense to dictate who speaks on campus**, it creates a chilling effect where members of the **political minority** no longer feel comfortable expressing their opinions. Diversity of opinion on campus benefits everyone, and the best way to ensure that diversity exists is to make sure people from many different perspectives feel comfortable speaking up. Essentially, **by banning speakers who do not espouse liberal views**, the university would be allowing the freedom of speech of one group to override the freedom of speech of another.

Note. Participants did not see text in bold

Dependent measures.

Expressed opinion. Each email was evaluated based on the opinion expressed by the participant (*Expressed opinion*). Emails were classified as *Pro ban* (*n* = 54/155) if they expressed support for the view that the protection of marginalized groups should override freedom of speech. Emails were classified as *Oppose ban* if they expressed the view that freedom of speech should be valued above the protection of minority groups. Emails were still classified as *Oppose ban* (*n* = 69/155) if they referenced an exception for inciting violence (e.g., *"A speaker should not be allowed to come to a campus if they have a history of inciting violence with their direct words*"). This exception was made because the Supreme Court has ruled that speech that incites violence is not protected under the First Amendment, so if participants did not think the university should protect this speech, they could still value freedom of speech over other values. The remaining emails were categorized as *Noncommittal* (*n* = 32/155). Emails were placed in this category if they either a) they discussed both sides of the argument (e.g., *"There is a very fine line between hate speech and free speech and that line needs to be considered when deciding who should get to speak on a college campus")*, b) explicitly expressed uncertainty (e.g., "*This is an interesting topic that I feel torn on because I've heard many compelling arguments from both sides"*), or c) qualified their opinions such that their positions were unclear (e.g., "*I do not think that certain speakers should be banned from having discussions on campus as long as they are academic and respectful"*).

Engagement in conversation. As a dependent measure to test H4 and H5, participants were asked how engaged they were during the conversation. This was measured

using five 7-point Likert type items (e.g., *I felt engaged in the conversation*; *M* = 5.84, *SD* = .97; α = .84).

Linguistic norms. Linguistic content of PC and other norms were determined using the Meaning Extraction Helper (MEH) (Boyd, 2017a) which aids in automated meaning extraction. When a set of texts are analyzed using MEH, the program returns a list of the most common n-grams, for example unigrams (one word) and bigrams (two words that occur consecutively) in the text, as well as the frequency of that n-gram in each text. Initial exploratory analyses indicated too few bigrams occurred at a sufficiently high frequency for analysis. There were only 27 bigrams that were used by more than 20 participants, whereas there were 90 unigrams that were used frequently. Therefore, for the purposes of this study, unigrams were analyzed.

Integrative complexity. Though the measurement of IC can involve laborious manual coding, it is also possible to automate its measurement using a CATA proxy measure. The most widely used automated method for evaluating IC is through the software Linguistic Inquiry Word Count (LIWC). The 2015 version of LIWC has more than 80 word categories that the software uses to determine the frequency count of words in a text that fall into any given category. Researchers have previously used LIWC to construct a successful proxy measure for IC (Abe, 2011; Brundidge et al., 2014). Because IC involves differentiation between ideas and their integration, it is possible to search for words that are indicative of contrasts and combining ideas in order to quantify the extent to which a text compares different views and integrates them. Although LIWC is a relatively crude measure (Conway, Conway, Gornick et al., 2014), it has nonetheless proven to be a valid automated indicator of IC. Specifically, the LIWC measure of IC uses exclusive words, negations, tentative words,

and conjunctions (Brudidge et al., 2014; Graesser, McNamara, Louwerese et al., 2004) (see Table 5). One change had to be made from the previously validated measure. Abe (2011) and Brundidge et al. (2014) used LIWC 2007. In the current study, the updated LIWC 2015 was used. According to Pennebaker, Boyd, Jordan et al. (2015) the category *exclusive* was replaced by a conceptually similar category, *differentiation* (e.g., hasn't, else). Therefore, *differentiation* was analyzed instead of *exclusive*. These items also create a reliable scale[10] (e.g., Brundidge et al., 2014). Two composite measures of integrative complexity were created, one for the email portion of the study (M = 4.77, SD = 1.11; α = .54) and one for the chat portion (M = 4.97, SD = 1.20; α = .63).

Covariates and moderators.

Endorsement of freedom of expression. In order to establish the extent to which participants were conforming to their partner (to test H1), participants' opinions on the topic prior to exposure to the experimental condition was measured. Therefore *Endorsement of freedom of expression* was measured using a 5-point six item scale from the 2005 and 2006 State of the First Amendment surveys. An exploratory factor analysis indicated that the items factor into two scales. The first scale, *Anti-Government Speech,* was made of two items (*Newspapers should be allowed to freely criticize the U.S. military about its strategy and performance; Even during wartime, the press should be allowed to publish stories that criticize the actions of the government*; M = 3.73, SD = 1.08; α = .85). The second scale, *Offensive Speech,* was made up of three items (*Musicians should be allowed to sing songs*

[10]Though these reliability scores appear low compared to more traditional numeric scales, this is normal for natural language data. Because there is much more variability in the words people used than in scores they might report on more traditional scales (Chung & Pennebaker, 2008), standard reliability criteria are not appropriate for evaluated natural language measures.

with lyrics that others might find offensive; People should be allowed to say things in public that might be offensive to racial groups; People should be allowed to say things in public that might be offensive to religious groups; *M* = 2.78, *SD* = 1.12; α = .85). Participants' score on *Offensive Speech* was presumed to measure their opinion regarding whether or not speakers should be banned from campus prior to exposure to the experimental manipulation.

Table 5.
Integrative Complexity Measure from LIWC Categories

IC Dimension	Category	Examples
Differentiation: words that make distinctions between ideas		
	Differentiation	but, either, without, whether
	Negations	not, none, never, cannot
	Tentative words	partly, overall, depending
Integration: words that join ideas together		
	Conjunctions	also, because, however

See Abe, 2011;2012; Brundidge et al., 2014

Political position. Political position was measured by two 7-point semantic differential items (1 *Strongly in favor of Democrats;* 7 *Strongly in favor or Republicans;* 1 *Extremely liberal,* 7 *Extremely conservative*; *M* = 2.82, *SD* = 1.27; α = .87).

Willingness to self-censor. WTSC was measured using Hayes et al.'s (2005) eight item scale (*M* = 3.02, *SD* = .78; α = .83).

Fear of isolation. FSI was measured using a five item scale (Hayes, et al., 2011) (*M* = 3.54, *SD* = .87; α = .83).

Manipulation check.

Perceptions of confederate opinion. After the completion of the text-chat, participants were asked to report what they believed their partner's political position was. The item was a 7-point semantic differential (1 *Very liberal*, 7 *Very conservative*; $M = 2.97$, $SD = 1.63$).

Results

Several methods were used to test the hypotheses and answer the research questions. Where appropriate, GLM, logistic regression, and multiple regression were used. See Table 6 for bivariate correlations between moderator and covariate measures. The linguistic analyses were conducted using LIWC and MEH.

Table 6.
Correlations Between Covariates and Moderators

	1	2	3	4
Enorsement of freedom of expression (*Offensive speech)*	1.00	.16*	.08	-.12
Political ideology		1.00	.30***	-.04
WTSC			1.00	.25***
FSI				1.00

* $p < .05$, ** $p < .01$, *** $p < .001$

Manipulation Checks

In order to ensure that participants interpreted the email and the chat in each condition in the expected way, a one-way ANOVA was performed, with condition as the grouping variable, predicting perceptions of the confederate's political ideology. The results indicate that participants in the Non-PC condition ($M = 3.67$, $SD = 1.75$) thought their partner was significantly more conservative than those in the PC condition ($M = 2.24$, $SD = 1.09$), $F(1, 154) = 37.25$, $p < .001$. The manipulation was successful.

H1: Self-Censorship Through Conformity

The first hypothesis predicted that the more conservative a participant was, the more likely they would be to conform to the confederate in the PC condition, but not the non-PC condition. This hypothesis was tested using multinomial logistic regression. Participants reported their opinions on the issue of whether or not speakers who offend minority groups should be censored (as measured by the *Offensive Speech* subscale of the freedom of expression scale) prior to coming to the lab. Political ideology and *Offensive speech* were weakly correlated (r = .16, p = .05), such that the more conservative a participant was, the more acceptable they thought it was to allow offensive speech. However, because this was a weak correlation, political ideology was not a precise enough proxy for opinions on freedom of expression, and *Offensive speech* was included in the model as a control variable. The dependent measure was participants' expressed opinion from the email exchange. This hypothesis was tested by entering condition (PC = 0, non-PC = 1) and political ideology (mean centered) into a multinomial logistical regression (and specifying their interaction term) as independent variables controlling for *Offensive speech*, and predicting participants' expressed opinion in the email portion of the study (-1 *oppose ban*, 0 *noncommittal*, 1 *pro ban*). The regression analysis indicated main effects of *Offensive speech*, $\chi^2(2, N=150) = 24.39$, $p < .001$, such that those who were more if favor of allowing offensive speech were more likely to be classified as *oppose ban*, B = .98, SE = .23, p = .02, odds ratio = 2.06, than *pro ban*. *Pro ban* did not differ from *noncommittal.* There was also a main effect of condition, $\chi^2(2, N=150) = 14.05$, p = .001. Those in the PC condition were less likely to express that they were in favor of banning speakers (vs. opposed banning them), B = 1.68, SE = .50, p = .001, odd ratio = 5.36, or express a noncommittal response (vs. oppose ban), B = 1.34, SE = .52, p = .01, odd ratio = 3.82. Condition was not a significant predictor of

differences between those who wrote noncommittal responses compared to those who were in favor of banning, B = -.34, SE = .51, p = .51, odd ratio = .71. The predicted interaction between condition and political ideology was nonsignificant, $\chi^2(2, N=150) = .87$, p = .64. The results indicate that even when controlling for participants' position on the subject of censorship prior to coming to the lab, those in the PC condition were not more likely to conform to the confederate's position. Therefore, there is no support for H1.

RQ1: Linguistic Content

The first research question asked what communication content would emerge from the email messages and chats. This research question was answered using the Meaning Extraction Helper (MEH: Boyd, 2017a). The unit of analysis was all of the text from the email (for the email portion of the study), and the combined text of the whole chat for each participant. When using MEH, the user inputs a corpus text files (i.e., each participant has their own file), and then the program creates a list of commonly used words across the corpus. The program has an extensive list of built in stop words (i.e., words to skip over in analyses) (e.g., *as, there*) and a list of word conversions, some of which account for common spelling errors or differences between US and British English (e.g., *colour* changes to *color*), as well as conversions for different forms of the same word (e.g., *disgusting* gets classified as *disgust*). Several word conversions that are not built into the software but are relevant to the context of the study were added manually. In addition, after conducting an initial analysis, I added several stop words that were common enough that they appeared in the list of highly frequent words, but did not provide additional meaning to the analysis. The following stop words were added to MEH's default list: *said,* which was commonly used in the phrase *with that being said* and *long* which was often used in the phrase *as long as*. After

all user-entered specifications have been made, MEH reads all text files in the specified corpus and creates a list of frequently used words. In this study, a 10% cutoff point was used, meaning words were only included if they occurred in at least 10% of texts (top 10% of emails and top 10% of chats respectively). Then a dataset is created that includes each observation (i.e., each participant), with a binary variable that indicates whether or not each participant used the word.

The results from the MEM for the email text and the chat text were analyzed separately, but both were analyzed using the same procedure. A PCA was conducted on the binary outputs of the MEM looking for co-occurrences of words. Though an Eigenvalue of one is often used as the cut-off for determining the number of components in a PCA, this can often lead to over or under-factoring (Brown, 2006). In this case, with the number of words that reached above the 10% threshold in the emails (82), as well as with the small sample size in comparison with many other studies that examine the nature language (e.g., LeFebvre et al., 2015), a cutoff of an Eigenvalue of one would have led to over-factoring (32 components from the emails; 34 components from the chat text). Therefore, the number of factors was determined by using Scree plots, which plots Eigenvalues from the PCA. In this test, the researcher looks for the 'bend' in the Scree plot, or looks for the place where each additional component (hereafter called 'factor') is no longer contributing more variance to the factor solution (Brown, 2006). Once the appropriate number of factors was determined from the Scree plot, the PCA was repeated, but the number of factors to extract was specified, and then the component structure was rotated using a Promax rotation, an oblique rotation method that allows factors to be correlated with one another. Last, using a cutoff factor score

of .30[11], factors were created. Items that did not load onto a factor, as well as items that loaded onto more than one factor, were not included in further analyses. Words that had a negative loading on their factor, indicating that the presence of that word meant a text was less likely to be included in that factor, were reversed prior to building composite measures.

Once factors are created, the researcher then interprets the words that make up the factor in order to determine the category that the words represent. Last, the researcher can then inductively group these factors into larger categories or themes, with the factors acting as subcategories (LeFebvre et al., 2015). These categories are not valenced, but rather are meant to represent themes or topics that appear in texts, based on words that frequently co-occur. Therefore, though it is possible to infer that a topic was present in a text simply by analyzing word usage, it is not always possible to infer the position that the writer took on that topic.

Category creation. Following the PCA of the email texts, the scree plot indicated that a seven factor solution was appropriate. The seven factors accounted for 25.74% of variance. Eight words (*speak*, *student*, *ban*, *listen*, *group*, *right*, *view*, and *privileged*) loaded onto more than one factor. These items were dropped, and the PC was conducted again before forming final factors and composite variables. Each of these factors was then made into a composite variable by reverse-coding any indicators that negatively loaded onto their factor, and then taking a mean score. Because each indicator is binary, that meant that composite scores could range from zero (if the participant did not use any of the words from

[11]This loading cut-off was chosen because it optimized the balance between minimizing cross-loading. Though the sample is relatively small for the number of categories extracted through the PCA, larger factor structures can often be stable with small sample sizes if each factor has more than four indicators (Wolf, Harrington, Clark, et al., 2013). MEM is flexible and allows the researcher to maximize the interpretability of their findings (Chung &

that factor) to one (if the participant used all of the words from that factor), so the composite indicates a ratio of words in the factor that participant used against those they did not use. Essentially, a higher score indicates that a participant used more of the words from that factor, indicating that the topic that factor represents was more of a focus in their email. The factors are: *Discrimination* (indicated by *homophobic, misogyny, racism*; *M* = .12, *SD* = .30), *Safe space* (indicated by *safe, marginalize, learn* (reverse coded), *perspective* (reverse coded)*; M* = .66, *SD* = .22), *Deliberation* (indicated by *understand, side, voice, space, hear, share, opinion, disagree, oppose, belief, negative* (reverse coded); *M* = .28, *SD* = .18), *Presentation* (indicated by *topic, controversial, issue, give, education, attend*; *M* = .28, *SD* = .18), *Diversity of opinion* (indicated by *protect, liberal, minority, college* (reverse coded), *campus* (reverse coded), *create* (reverse coded), *bring* (reverse coded); *M* = .39, *SD* = .19), *Campus life* (indicated by *agree, school, speaker, comfortable, free* (reverse coded), *speech* (reverse coded); *M* = .38, *SD* = .23), and *Role of university* (indicated by *allow, fact, hate, political, society, attack, university, benefit*; *M* = .25, *SD* = .18). See Appendix B for words that make up each factor, factor loadings, and examples of those words in the email texts.

Last, the researcher can group these factors into larger overarching categories (LeFebvre et al., 2015). I examined the factors and texts that used the words that made up each factor, and inductively created categories that represent the topics of the email texts. In answering RQ1, the linguistic content of the participants' emails indicate that participants were focused on three main topics. The first category, PC, was made up of *Discrimination* and *Safe space.* Emails that included these factors discussed the protection of minority groups, and generally advocated positions consistent with liberal PC ideology. The second

---

Pennebaker, 2008), and .30 cutoff allowed for the simplest interpretation.

category, Public Deliberation, was made up of three factors: *Deliberation*, *Presentation*, and *Diversity of opinion*. Responses that included words from these factors tended to discuss issues related to the free exchange of ideas. The last category, University was made up of *Campus life* and *Role of university*, and these responses discussed the impact speakers might have on student life, as well as what responsibility the university has to either protect minority groups or protect freedom of speech. See Table 7 for interpretation of each category.

Table 7
Descriptions of Factors from Emails

Category	Factor	Words	Description
PC			
	Discrimination	homophobic, misogyny, racism	Explicit references to discrimination of marginalized groups
	Safe space	safe, marginalize, learn (reverse), perspective (reverse)	Highlighting the importance of protecting marginalized groups
Public Deliberation			
	Deliberation	understand, side, voice, space, hear, share, opinion, disagree, oppose, belief, negative (reverse)	Discussions of the elements of public deliberation.
	Presentation	topic, controversial, issue, give, education, attend	A focus on what is appropriate for public presentation.
	Diversity of opinion	protect, liberal, minority, college (reverse), campus (reverse), create (reverse), bring (reverse)	Highlighting the importance of the university is exposing students to a variety of viewpoints
University			
	Campus life	agree, school, speaker, comfortable, free (reverse), speech (reverse)	Focusing on the ways in which bringing/banning these speakers affects the overall college experience for students
	Role of university	allow, fact, hate, political, society, attack, university, benefit	Focusing on what the purpose of a university is. Present both responses advocating freedom of speech and those advocating protecting minority groups.

H3: Integrative complexity

The third hypothesis predicted that the more liberal a participant was, the lower they would be in IC, and that there would an interaction between political ideology and condition, such that liberal participants would show lower IC in the PC condition than in the non-PC

condition. In order to test this hypothesis, a 2 (Condition: PC/non-PC) by political ideology GLM was performed, predicting IC in the email and chat texts as two separate models.

The results indicate that condition was a significant predictor of IC, $F(1, 146) = .7.20$, $p = .008$, $\eta_p^2 = .047$, such that those in the non-PC condition exhibited higher IC ($M = 5.04$) than those in the PC condition ($M = 4.56$). There was also a significant effect of political ideology, $F(1, 146) = 5.02$, $p = .027$, $\eta_p^2 = .033$, such that the more conservative a person was, the higher they were in IC (B = .21, $t = 2.11$, $p = .036$). The predicted interaction between condition and political ideology was nonsignificant, $F(1, 146) = .77$, $p = .38$, $\eta_p^2 = .005$. These results indicate a main effect of political ideology, but also include an unpredicted main effect of condition, so H3 is partially supported.

H4 & H5: Individual Difference Variables

The fourth and fifth hypotheses predicted a three way interaction between condition, political ideology, and (H4) WTSC and (H5) FSI, predicting participant self-reported engagement in the conversation, such that in the PC condition, for more conservative participants, the higher they were in these individual difference variables, the less engaged they would report being in the conversation. H4 was tested using a 2(Condition: PC/non-PC) by political ideology (mean centered) by individual difference variable (substituting in WTSC and FSI; mean centered) GLM predicting self-reported engagement in the conversation.

WTSC. In order to test H4, WTSC (mean centered) was included in the model as the individual difference variable. The results indicate a main effect of WTSC, $F(1, 142) = 7.80$, $p = .006$, $\eta_p^2 = .052$, such that the higher a participant was in WTSC, the less engaged they were in the conversation ($\beta = -0.24$, $t = -2.79$, $p = .006$). All other main and interaction effects

were nonsignificant (all *F*s < 2.51). Because there was only a main effect of WTSC but no interaction, H4 is not supported.

FSI. In order to test H5, FSI was substituted for WTSC in the GLM, predicting participant self-reported engagement in the conversation. When FSI was substituted into the equation, all main and interaction effects were nonsignificant (all *Fs* < 1.28), including the main effect of FSI, $F(1, 141) = .03$, $p = .85$, $\eta_p^2 < .001$. There is no support for H5.

Discussion

The purposes of Study 3 were to replicate the findings of Study 2 with a behavioral dependent variable, to analyze the linguistic content and themes of conversations about PC salient topics, to examine the relationship between PC and integrative complexity (IC), and to test the potential moderating effects of the individual difference variables WTSC and FSI. In serving this first purpose of replicating the findings from Study 2, H1 tested the extent to which participants in the PC condition conformed to their partner's opinion more than those in the non-PC condition. The results indicate that participants did not explicitly conform to their partner's opinion the in PC condition. However, due to several methodological issues, the extent to which this study can answer questions regarding the effects of PC on self-censorship is limited. In serving the second purpose of this study, the language from the email portion of the study indicates that participants were focused on topics related to protecting marginalized groups, what role the university should play in protecting these groups versus supporting free speech, and the importance of public deliberation. In testing the relationship between PC and IC, the results indicate that participants in the PC condition exhibited lower IC than those in the non-PC, and more liberal participants exhibited lower IC than more conservative participants. Last, in testing the individual difference variables, the

higher a participant was in WTSC, the less engaged they reported being the in the conversation. There were no effects of FSI. Despite several methodological flaws that make it difficult for these data to speak to the relationship between PC and self-censorship, the results pertaining to language content, accommodation, and IC do provide insight into the effects of PC on public deliberation.

Forms of Self-Censorship

There was no evidence that participants conformed as a form of self-censorship. Those who self-reported in the pre-survey that they were either in favor of or opposed to banning public presentations by potentially offensive speakers were also likely to express that opinion in their email to their partner. Though this does not support the original hypothesis, this is likely a reflection of the fact that opinion expression and self-censorship are not as simple as conforming to a communication partner's opinion. For example, in Hayes' (2007) study where participants were asked how they would act in a hypothetical scenario, presenting a false opinion (i.e., conforming to one's hypothetical communication partner) was the least cited strategy participants reported that they would engage in. They were more likely to report that they would reflect back any questions that were asked of them, express uncertainty, or feign indifference. In light of Hayes' finding, the results of the current study should not be surprising.

Due to a methodological flaw in the study, it is difficult to draw conclusions about self-censorship. Participants were never asked to self-report on their opinion expression or self-censorship during the conversation. A self-report measure of participants' perceptions of their communication would have provided more insight. In order to overcome this methodological flaw, a replication study that has such a self-report measure will be

necessary, and the texts created through this study will be manually coded. Human coders can examine opinion expression/self-censorship by coding for the avoidance strategies outlined by Hayes (e.g., reflecting questions, expressing indifference, making a joke) and other interpretations of the conversations (e.g., adding qualifiers to argument, expressing a lack of knowledge). Through manual coding and a replication including a measure self-censorship, self-report data can be used to validate linguistic predictors of self-censorship.

Linguistic Content of Norms

Though much research has examined the relationship between political position and language use in news (e.g., Haertl & McCarthy, 2010; Terwilleger, McCarthy, & Lamkin, 2011), blogs (Brundidge et al., 2014), and political speeches (Benoit & Laver, 2006; Slapin & Proksch, 2008), to my knowledge, this is the first study to examine the linguistic patterns of interpersonal discussions about PC relevant topics. The linguistic content of the emails indicate that when discussing a PC relevant topic, a variety of themes emerged, including discrimination, protecting minority groups, university life, diversity of opinion, and freedom of expression. These topics mirror many of topics which have been historically associated with PC. For example, the PC debates of the 1990s were largely about what should be done to reduce racial and gender discrimination, how this should be balanced against freedom of speech, and specifically how the values of reducing discrimination and freedom of expression should be balanced on college campuses (Berman, 1992; Hughes, 2010). The MEM results indicate that many of the arguments made in early years of the PC movement at universities continue to occur today. However, it is important to keep these findings in perspective. In order to serve the purposes of answering the overarching research questions of this dissertation, it was necessary that participants be primed by reading the confederate's

opinion first. However, in order to draw more generalizable conclusions about the nature of the linguistic content of PC norms and the topics that participants think are important, future studies should examine texts that were written about PC topics without being primed by a confederate's position first. In doing this, it will be possible to gain a more accurate insight into the nature and prevalence of PC linguistic norms, as well as themes that emerge from conversations about PC relevant topics.

Integrative Complexity

This study has made contributions to our understanding of public deliberation. When true public deliberation occurs (which need not be "public" in the traditional sense of the word, as it can occur in any conversation between citizens; Carpini et al., 2005), those involved not only acknowledge different points of view, but they also consider the costs and benefits of these views as they revise their own opinions (Jackman & Sniderman, 2006). IC, which accounts for both the differentiation and integration of ideas, is a way to measure the extent to which people are upholding this ideal of deliberation. The results show that in the email portion of the study, participants in the PC condition exhibited lower IC than in the non-PC condition. Furthermore, more liberal participants exhibited lower IC than more conservative participants. Research has generally found that liberals display higher IC in their arguments (Tetlock, 1983). The explanation generally given is that conservative ideology has a tendency to favor good/bad explanations for problems, rather than understanding the complexities of issues (Jost, Glaser, Krulanski et al., 2003). However, as Conway et al. (2016) found, complex thinking is context dependent, and there are political issues on which liberals have less complex thinking than conservatives. Robinson and Reid (2016a) argue that

PC, which is commonly associated with liberal ideology, is often advocated in a similar good/bad manner.

The issue participants discussed in this study, whether or not to ban controversial speakers from campus, as well as most other political issues, are complex and require consideration of several different values in order to make informed decisions. College campuses are the ideal time for young people to learn how to consider these complexities. If PC language primes ideology and identity politics, and ideological commitment is associated with less complex thinking, then it is important for instructors to consider how they teach about controversial issues that relate to students' identities. Though there are many pedagogical strategies that instructors can use (e.g., playing devil's advocate, assigning students to roles in debate; Avery & Steingard, 2008), the results of this study, along with results of Study 2, bring to light another possibility: if instructors are to effectively encourage rational deliberation, they should be aware of their language use and how it may prime ideology.

Individual Differences

A contribution of this study is that it is one of the first to examine the effects of WTSC and FSI on self-reported engagement in conversation that had just occurred, rather than relying on participant's self-reports on how they would act in a hypothetical scenario. First, the results show that participants who have a tendency to self-censor (i.e., high in WTSC) reported being less engaged in the conversation. However, the lack of significant interaction between the individual difference variable, condition, and political ideology calls into question the claims of Hayes and colleagues that WTSC affects those in the opinion minority but not those in the opinion majority (e.g., Hayes 2007). FSI, on the other hand, was

not predictive of engagement in the conversation. One reason for the difference in how these variables predict engagement in the conversation may be that although both measures come out of research on SoS and tend to be strongly correlated (Hayes et al., 2011), FSI is focused on the motive one might have for staying silent (avoiding social sanctions) and WTSC purposefully does not account for motive (Hayes et al., 2005). Hayes et al. (2011) point out that these variables measure different constructs. Because the research assistants were trained to give similar responses to everyone, and none of those responses included any type of social sanction (e.g., no insults, no cursing), it may be that as participants moved through the conversation, they realized that there was a low risk of incurring social sanctions and did not fear social isolation. Overall, more research is needed in order to understand how and why FSI and WTSC may be predictive of self-censorship in different contexts.

Limitations

As mentioned above, in order to make more conclusive claims about the effect of PC on self-censorship, a replication study that includes an explicit self-report measure of self-censorship should be conducted. Furthermore, the texts generated from this study should be analyzed by a group of coders who can examine nuances of opinion expression/self-censorship that cannot be captured by CATA.

Furthermore, the experimental manipulation confounded language use with political position. In the PC condition, the confederate used PC normative language to advocate a PC consistent position. In the non-PC condition, the confederate used a non-PC language in the email, in order to advocate a PC-inconsistent position. Therefore, unlike Study 2, Study 3 does not test the ways in the which language primes ideology and is subsequently related to self-censorship, but instead can only speak to the ways in which participants communicated

with a partner who advocated a position consistent or inconsistent with PC. With that being said, the data still indicate that participants did communicate with more PC language and lower IC in the PC condition, which still provides insight into the relationship between PC and public deliberation.

There are two other limitations that must be addressed. First, the sample used in this study was 87.1% female. Past research has found that men and women use language differently. For example, compared to men, women tend to ask more questions, express more uncertainty, and use more negations (for review, see Mulac, Bradac, & Gibbons, 2001). This is meaningful because some of these word categories are also included in the measure of IC (e.g., negations). However, because research has not found meaningful sex differences in IC (Suelfeld, Tetlock, & Streufert, 1992) the sex-imbalance in this study should not have been enough to change the results. Second, is the effect of using confederates as the conversation partner. Though using confederates provides experimental control, their instructions also affected the conversations. Because the behavior of interest was participants' natural language, the research assistants were instructed to use comments from the pre-written list of responses and to ask questions. Because the confederates were instructed to give the participants as much opportunity as possible to generate text, this may have made the confederate seem more open and willing to listen than most conversation partners would. If the study had been conducted between two participants, rather than one participant and a confederate, the results likely would have included more variability in self-censorship and opinion expression, as there would have been more variety in the combinations of people (e.g., one dominant person/one self-censor, more extreme political ideologies, more natural

lulls in conversation, etc.). Future studies should examine how two naïve participants engage in these conversations.

Despite these limitations, this study provides an important first step in the examination of the effects of PC norms in actual conversations, and how that ultimately affects the complexity of the arguments used in debates about important topics. Future studies should be conducted in order to overcome the limitations of this study.

Chapter VI: Conclusion

This dissertation has made contributions to our understanding of self-censorship, rational deliberation, and the role that political correctness plays in these phenomena. In what follows, I summarize the key findings, discuss theoretical and practical implications, as well as discuss limitations. I conclude by proposing a model of the process of how intergroup conflict can lead to self-censorship through increases in identity salience and PC ideologies, and in turn, decreased public deliberation.

Overview of Findings

Study 1 tested the extent to which participants' political ideology was related to their perceptions of victimization on campus, and how that in turn was related to who they thought was deserving of censorship. The findings indicated that the more liberal a participant was, the more likely they were to think that members of racial minorities, sexual minorities, Muslims, and women are the victims of discrimination. However, the more conservative a participant was, the more likely they were to perceive that Whites, Christians, and men were the victims of discrimination. Similarly, ideology was also predictive of who participants thought deserved to be censored, such that liberals were more in favor of banning speakers who advocate conservative positions and conservatives were more in favor of banning speakers who advocate liberal positions. Additional analyses indicated that People of Color and women were not more likely to perceive victimization among marginalized groups, nor were they more likely to support the censorship of conservative speakers, indicating that political ideology, rather than membership in a marginalized group, is a better predictor. Perceptions of victimization were not necessarily used as a justification for censorship (i.e., perceptions of victimization did not mediate the relationship between political ideology and

support for censorship), though this may have been due to methodological issues (wording of some items, in particular). Overall, the findings of Study 1 indicate that perceptions of victimization and opinions about who should be allowed to express their opinions on campus are often determined on ideological grounds that are rooted in identity politics.

Study 2 tested whether PC ideology is made salient by primes of PC normative language, and how that, in turn, was related to self-censorship. Specifically, Study 2 showed that both liberals and conservatives interpreted PC language used by an instructor in a classroom setting as discouraging conservatives from expressing their opinions, indicating the PC normative language acted as a cue to the speaker's ideological commitment. Furthermore, more conservative participants were less likely to report that they would express their opinions when in the PC prime condition than when in the non-PC condition. Study 2 also showed that both People of Color and White students recognized the language used by the instructor in the PC condition, but not in the non-PC condition, as encouraging expression from racial minorities. Furthermore, in the PC condition, White students were less likely to report that they would share their opinions than were People of Color, but People of Color were not similarly discouraged by the non-PC condition.

Study 3 tested the extent to which the results of Study 2 would be replicated in an actual conversation about a PC relevant topic (whether or not to allow potentially offensive speakers to come to campus). Study 3 also examined how people engage in debates about these topics through analyzing linguistic trends and integrative complexity. A methodological error (the omission of a self-report measure of self-censorship) prevented me from being able to serve this first purpose of replicating Study 2. This flaw will be remedied in replication studies. Nonetheless, in serving the second purpose of Study 3, using the meaning extraction

method, the linguistic content of the email portion of the study indicated that participants were focused on PC relevant issues (specifically, discrimination and having safe spaces), public deliberation (specifically, the extent to which allowing potentially offensive speakers can enhance or hurt public deliberation), and how those in the university administration should balance the need to protect minority students against the value of freedom of expression. In serving the third purpose of the study, a measure of integrative complexity (a proxy for rational deliberation) indicated that participants in the PC condition (i.e., an email conversation where a confederate communication partner expressed a PC consistent position supporting banning a potentially offensive speaker while using PC code words) made arguments that were less integratively complex than those who were in a non-PC condition (i.e., in which the confederate opposed banning potentially offensive speakers). Furthermore, the more liberal participants made arguments lower in integrative complexity than more conservative participants.

Theoretical Contributions

Expanding on argument made by Robinson and Reid (2016a), in this dissertation I made initial steps in testing the extent to which PC may be used as a tool to control discourse. This perspective has the potential to make contributions to research on social identity, self-censorship, and public deliberation.

Social Identity Approach

According to SIT (Tajfel & Turner, 1979), people try to maintain a positive social identity by having a positive evaluation of the social groups to which they belong. If they are not able to maintain a positive social identity (because their group has relatively low status or because they have a low status position within the group), they are likely to engage in

strategies to change their position. People can engage in social mobility (i.e., changing groups) or social creativity (i.e., changing the way they think about their group memberships) if direct confrontation is undesirable. When these two strategies are not possible or not desirable, groups are more likely to engage in social competition, which involves direct conflict with rival outgroups. In line with SIT, Tajfel (1982) posits that stereotypes about outgroups serve the ideological function of providing groups' explanations for their position in society, as well as allowing them to justify actions against outgroup. In Chapter II I argued that PC is a specific form of social competition where groups use claims about threats posed by the outgroup in order to control their political rivals through the control of discourse. The results of Study 1 support this explanation of PC. The findings indicate that political ideology is predictive both of perceptions of the victimization of coalitional partners (members of racial, sexual, and religious minorities and women for liberals, and Whites, Christians, and men for conservatives), and the desire to censor outgroup speakers. Specifically, liberals desired to censor speakers who advocated a conservative position and conservatives desired to censor speakers who advocated a liberal position. Moreover, Study 1 found that membership in a marginalized group (at least for racial groups and sex), was not predictive of perceptions of victimhood or support for censorship above and beyond the predictive power political ideology. This further supports the claim that PC is at least partially about the control of political rivals' communication.

Furthermore, as was found in Study 2, simply priming liberal PC ideology through the use of normative language was enough for participants to categorize the teacher as someone who was encouraging minority groups to share their opinion. This is consistent with self-categorization theory (Turner, Hogg, Oakes et al., 1987), which posits that group

prototypes about attitudes and behavior act as a way for people to distinguish who belongs to which group. Study 2 found that the use of prototypical PC code words was enough to prime this categorization process. Furthermore, SCT posits that these prototypes not only aid in social categorization, but also aid in managing intragroup processes. To this end, Study 2 not only found that language primes allowed people to engage in social categorization, but also it was also predictive of participants' self-reports of how they would behave. For those in the PC condition, the more conservative a participant was, the more likely they were to report that they would self-censor. Considering the small number of actual conservatives in the sample, it was also many political moderates and even moderate-liberals who reported that they would self-censor under the PC prime. This likely indicates the PC prime made relevant prototypes salient, thereby indicating what would be considered normative attitudes and behaviors (i.e., expressing PC consistent views). For those whose social identity is such that they value being a part of a classroom "community," regardless of whether or not they held prototypical attitudes, there would be an incentive to engage in prototypical behaviors (meaning a view that could be interpreted as being non-PC).

These findings support the argument made by Robinson and Reid (2016a) that adherence to PC is often due to forms of social influence other than persuasion. According to Turner (2005), people may change their behavior for one of three reasons. First, they may be persuaded to do so. In the case of PC, there are many people who genuinely embrace not just the ideals behind liberal PC ideology (i.e., racial and sexual equality), but also its more specific prescriptions (e.g., normative language, explanations for why inequality exists, ideas about how equality can be gained). However, there may be others who subscribe to the ideals of PC ideology but not all of its corresponding norms who may nevertheless abide by these

prescriptions because some higher authority has indicated that they should do so. Though this authority may dictate these prescriptions explicitly (for example, I once was the teaching assistant for a class that had to following statement in its syllabus for prescriptions on a class assignment: "use politically-correct content and language"), the results of Study 2 indicate that an authority need not be this explicit. In Study 2, participants inferred from a teacher, who is the authority in a classroom, what was normative and reported that they would act accordingly. Last, people may follow normative prescriptions because they are coerced to do so. This can occur through the threat of sanctions or through watching others be sanctioned for deviant behavior. This dissertation did not test the effects of sanctioning (see Limitations). However, overall, the results of this dissertation indicate that PC has at least some social influence at UCSB.

Self-Censorship

This dissertation has made contributions to our understanding self-censorship, and this contributes to research on the Spiral of Silence (Noelle-Neumann, 1974). According to the SoS, people have a quasi-statistical sense, and using this, they evaluate the opinion climate and make decisions about expressing themselves. Though research on the false consensus effect finds that people tend to overestimate the extent to which others agree with them (Mullen, Atkins, Champion et al., 1985; Ross, Greene, & House, 1977), research on the SoS has not explained how it is that a person would determine that others *disagree* with them. As Hayes, Matthes, and Eveland (2011) point out, the claim of SoS that people constantly seek out information in order to determine the opinion climate has gone untested. Rather than having participants use their quasi-statistical sense, tests of the SoS have relied on experimental manipulations where participants are either explicitly told what the majority

opinion is, or they are told to imagine that they are in the opinion minority or majority (see Glynn, Hayes, & Shanahan, 1997 for meta-analysis). In an attempt to test the concept of the quasi-statistical sense, Hayes et al. (2011) found that fear of social isolation, the supposed driving mechanism behind the spiral of silence, was correlated with the extent to which people pay attention to opinion polls. However, even in this test, Hayes et al. were examining the extent to which people pay attention to explicit cues about the opinion climate. Though there are rare occasions when people have access to this explicit information (e.g., polls during election season), these explicit cues are often too broad to be helpful in a specific context (e.g., national opinion polls are not helpful in determining the opinion climate in one's local community). More importantly, people rarely have access to this explicit information about the opinion climate and use other cues, such as those found in language, in making their assessment.

The results of Study 2 provide insight into how people use their quasi-statistical sense when evaluating less explicit information. Study 2 tested the extent to which language primes ideology, giving listeners an indication of what the dominant (or acceptable) opinion is. In this way, Study 2 was able to test the idea of the quasi-statistical sense. Study 2 found that language can prime ideology, indicating what the dominant ideology is within that particular environment. Specifically, Study 2 showed that both liberals and conservatives interpreted PC language used by an instructor in a classroom setting as discouraging conservatives from expressing their opinions, and more conservative participants reported that they would self-censor accordingly. This is an indication that people do use cues from the environment in order to determine whether or not to share their opinions. Future tests of SoS should use these more subtle cues in order to test the idea of the quasi-statistical sense.

Public Deliberation

Public deliberation is a necessary condition for democracy to exist (Carpini, Cook, & Jacobs, 2005), and requires that debaters express their viewpoints coherently and rationally (Habermas, 1984; 1989), be open to alternative viewpoints, and be willing to reconsider and modify their previously held beliefs (Jackman & Sniderman, 2006). A key component of the democratic process and rational deliberation is that people with different interests and ideas argue their case, feel that their views are represented, and talk *to* one another instead of *past* one another (Habermas, 1989). When groups are able to engage in this form of rational deliberation, even when dissatisfied with the results, both sides of an argument feel that the system is fair and that their interests have been represented (Dahlberg, 2005). In order for people from multiple perspectives to feel that their interests have been represented and the system is fair, as is required for public deliberation, groups with opposing positions must be able to express their views. The results from Study 1 indicate that participants' political ideology was predictive of who they thought should be banned from speaking on campus. The desire to censor speakers who espouse views counter to one's ideological position opposes the ideals of public deliberation. If public deliberation requires that deliberators be open to not just hearing the positions of others, but also weighing the merits of these different arguments, then it is difficult to embrace this ideal while still supporting the censorship of political opponents. In accordance with the arguments of many journalists and scholars who have expressed concern about the state of democratic processes at universities (e.g., Al-Gharbi & Haidt, 2017; Dreger, 2015; Kipnis, 2015; Loury, 1994; Lukianoff & Haidt, 2015; Schlosser, 2015), the results of Study 1 indicate that students' identity politics may be undermining public deliberation on campus.

Beyond support for censorship, Study 3 of this dissertation found evidence that when liberals discuss PC topics, they exhibit less complexity in their arguments than do conservatives. Integrative complexity is a measure of the extent to which a communicator is engaging in the ideals of public deliberation, as it measures the extent to which a person acknowledges different viewpoints and is able to weigh them against each other (Tetlock, 1983). The topic that participants discussed in Study 3 has valid arguments on both sides: protection of marginalized groups on one side and support of freedom of speech on the other. Because of the topic's relevance to the lives of university students (e.g, Al-Gharbi & Haidt), it is even more important that they be able to weigh these tradeoffs against each other. However, the data from Study 3 indicate that more liberal participants were not necessarily weighing these tradeoffs as they exhibited lower IC than did conservatives when discussing the topic. This is consistent with the findings from Conway et al. (2016), who found that liberals and conservatives both have issues about which they are less able to differentiate and integrate opposing viewpoints. For issues that are seen as fundamental to a particular ideology (abortion for conservatives, the protection of marginalized groups for liberals), they are less likely to see any grey area, and thus exhibit fewer of the communicative patterns necessary for engaging in true public deliberation.

PC is advocated in such a way that only strict adherence to the ideology is seen as "good", and a person who deviates from the ideology is classified as immoral (Robinson & Reid, 2016a). Research has found that liberals and conservatives tend to prioritize domains of morality differently. Liberals tend to be focused on fairness and avoiding harm. Conservatives also value fairness and avoiding harm, but give equal weight to ingroup loyalty, respect of authority, and sanctity (Graham, Haidt, & Nosek, 2009). Those who

advocate the PC consistent position that speakers who may be seen as offensive are likely to harm minority groups may be unlikely to see any room for deviation from this position as the issue is framed in terms of morality and avoiding harm, especially if that harm has the possibility of increasing inequality. Meaning, the topic of banning potentially offensive speakers is fundamental to liberals' sense of morality. Study 3 found that for the email portion of the study, participants in a PC condition exhibited less overall IC than those in the non-PC condition. It is possible that priming PC may have indicated to participants that they should not weight multiple sides of the argument, but rather that there could be only one "correct" side. More empirical research about the effects of priming norms on IC is needed, but overall, this study contributes to our understanding about how to facilitate public deliberation: if priming PC ideologies is associated with lower IC arguments, then deliberative democracy researchers should examine the ways in which more neutral language and environments can affect debate.

Additional Theoretical Tests for Future Research

Communication Competence. Spitzberg's (1991) model of communication competence posits that communication is competent when it is seen appropriate by the receiver and is effective in achieving the sender's goal. Sptizberg points out that communicative encounters do not occur in a vacuum. Each communicator has individual differences that affect the interaction (*individual system*), each situation has unique features that affect how the communicators are perceived (*episodic system*), and the interaction is also influenced by the relational history of the members of the dyad (*relational system*). Within the episodic system, it is possible to see the role that PC might play in perceptions of communication competence. For example, Spitzberg argues that receivers' perceptions of the

speaker's competence are likely a function of the receivers' expectations prior to entering the conversation. When a speaker meets or positively violates a receiver's expectations (e.g., when the receiver expects the interaction to be negative but it ends up being a positive interaction), they are more likely to be viewed as competent. Sptizberg references expectations that stem from our schemas of certain situations (e.g., what an interview should look like, how communication usually occurs in a doctor-patient interaction). There are likely other important factors that receivers consider when forming their expectations of an interaction, such has the group membership of the speaker.

But how do receivers' expectations affect interpretations of appropriateness? If a receiver expects a person to communicate inappropriately, perhaps because of their group membership, previous experience with that individual, or because they have some other motivation to find fault in the speaker's message, are they more likely to perceive that speaker as inappropriate? Research on motivated reasoning suggests that the answer is yes. If people are motivated to reach the conclusion that is consistent with their preferences (Ditto et al., 2009; Kunda, 1990), then they should be more likely to interpret communication as inappropriate, and that the speaker lacks competence, if that was their expectation prior to the interaction.

Why does this matter for the study of PC and its relationship to self-censorship? If communicators want to be viewed as competent, they not only have to consider their goals, but they must also consider what their communication partner will perceive as appropriate. This requires the speaker to guess what the receiver's expectations are for the interaction. If a speaker suspects that they will be viewed as inappropriate, either because of their group membership or because the topic is often viewed as inappropriate, they may decide to self-

censor. The results of Study 2 show that more conservative students were more likely to self-censor in a PC condition than in a non-PC condition. Though this study did not evaluate why participants were self-censoring (e.g., because they feared being perceived as inappropriate, they were trying to make space for other perspectives), it is possible that they did not know how to communicate competently in the PC condition. Future research on communication competence should examine both how receivers determine what is appropriate communication, as well as how message senders assess the likelihood that they will be perceived as appropriate when deciding whether to share their opinions.

Communication Accommodation Theory. Research testing CAT has found that appropriate communicative adjustments (which may be convergence, divergence, or maintenance, depending on the context) are positively evaluated. Though this generally occurs when accommodation is symmetrical (i.e., both partners are converging toward each other), there are times when asymmetrical communication (i.e., only one partner is converging) is seen as appropriate. For example, there are circumstances where only one interactant is able to accommodate due to language abilities, or where social roles dictate asymmetry (e.g., salespeople accommodate to customers more than the other way around) (for review, see Dragojevic, Gasiorek, & Giles, 2016). Dragojecvic et al. (2016) argue that asymmetrical communication is often perceived as complementary, rather than unequal, as these asymmetrically accommodative interactions are often rated positively by both interactants.

However, if, as Robinson and Reid (2016a) suggest, those who do not subscribe to PC often follow its normative prescriptions, then it is possible that there will be consistent asymmetrical convergence toward PC normative language. Though this dissertation did not

directly examine accommodative processes, some of the findings from this dissertation indicate that one way in which people may enact PC norms is through convergence. For example, Study 2 found that PC language does prime ideology, and more conservative participants in a PC (vs. non-PC) condition were more likely to report that they would self-censor. Furthermore, by manually coding the chats generated during data collection from Study 3, it will be possible to examine the extent to which participants converged toward the language of their communication partner, and whether this was more likely to happen in the PC condition. Though much CAT research examines the adjustment of language and nonverbal behavior, instances of avoiding certain topics or avoiding sharing one's opinion for the sake of avoiding conflict can also be considered convergence (Dragojevic et el., 2016; Soliz & Giles, 2014).

CAT research has generally examined individual interactions. These interactions often go more smoothly when interactants perceive that their partner has engaged in appropriate accommodation (even if asymmetrically) (see meta-analytical results in Soliz & Giles, 2014). However, when convergence is asymmetrical, not just for one interaction, but consistently across large numbers of people and interactions over an extended period of time, this has the distinct possibility of leading to negative effects. In fact, the prediction that consistent asymmetric convergence can have negative consequences is made by the theory of ethnolinguistic vitality (Giles, Bourhis, & Taylor, 1977; Giles & Johnson, 1981; Yagmur & Ehala, 2011). According to the theory, if a minority language and its speakers are of low status, are a small minority, and receive little institutional support, then those who speak that language will asymmetrically converge toward the dominant language, and the minority language is likely to die. It is possible that PC inconsistent positions can, perhaps, be seen as

low vitality positions: people are often sanctioned for expressing these positions, indicating that they are low status. This can lead to these positions being expressed less often, which can give the impression that they are minority opinions (in some cases these may actually be minority opinions, but in others it may be a misperception). It is possible that if PC norms have such authority that people are more likely to converge in their direction, an echo chamber can be created where only certain opinions are expressed. To my knowledge the predictions of ethnolinguistic vitality theory have not been used to examine opinion expression. Nonetheless, it is a useful framework for understanding how asymmetric communication on a community or societal level may cause a decrease in diversity of opinion, and should be explicitly tested in reference to diversity of opinion in future research.

Practical Implications

Debates about PC and related topics do not only occur among academics. Journalists, professors, and university administrators are also joining the conversation, asking questions such as: What role should university administration play in regulating guest speakers on campus? In ensuring viewpoint diversity, how should universities balance the need to encourage members from historically disadvantaged groups to express their opinions against the desire to encourage those in the political minority (which may be conservative students on some campuses, but may be liberal students on others) to also express their views? How should instructors most constructively facilitate discussions about sensitive issues in a way that exposes students to opposing viewpoints without making anyone feel threatened?

Ultimately, the purpose of this dissertation was not to provide conclusive answers to these questions, or at least not the extent of saying what universities *should* do. Each university has a different population of students with different needs. Each university has its

own mission and values. Each instructor has their own goals for their class and the environment they want to create in the classroom. I do not mean to tell university administrators or instructors what they *should* value. Instead, from a practical perspective, the purpose this dissertation is draw attention to some of the consequences of PC on university life, and to allow those involved in universities to decide how to balance these consequences against their own goals and values. In what follows, I outline some practical implications of the findings from this dissertation for university policy and classroom instructors.

University Policy

Though universities may differ in their specific goals, at a minimum, most claim to be committed to the education of their students and to diversity (Morphew & Hartely, 2006). However, universities may differ in exactly what they mean by *education* and *diversity*. Does education mean training students for their future careers, teaching them to be informed community members, or learning how to live in the "real world"? Does diversity mean having a university that is demographically diverse, creating an environment that has viewpoint-diversity, or a combination of the two? The ways in which a university answers these questions will determine how they balance the tradeoffs of PC against its potential effects on censorship, self-censorship, and public deliberation.

So how should universities use this information specifically in the context dealing with guest speakers on campus? One strategy is to ban such speakers. As Hentoff (1992) argued in the early days of the PC debate, "the cheapest, quickest way to demonstrate that [the university] cares is to appear to suppress racist, sexist, and homophobic speech," (p. 216). This appears to be the strategy employed by UC Berkeley in its deplatforming of

speakers when there was sufficient student/public opposition (e.g., Yiannoplous, Coulter, Dawkins, Maher). However, the results from Study 1 indicate that when people determine which speakers are too offensive to be considered for public presentation on campus, they often make these determinations based on their own political ideology. Essentially, it appears that students determine what should be banned based on what they agree with. When universities decide to ban speakers based on the demands of students, it not only has the possibility of silencing minority viewpoints and creating a chilling effect, but it also indicates to students that censorship is an appropriate way of dealing with dissent. This has the potential to harm anyone with an unpopular view, independent of the ideological position that view supports (Al-Gharbi & Haidt, 2017).

Given that universities are trying to balance multiple goals and values, including embracing various kinds of diversity (e.g., viewpoint-diversity that stems from coming from different backgrounds or different political perspectives), a more nuanced approach to dealing with guest speakers may provide benefits to campus life. In contrast to UC Berkeley, UC Santa Barbara has employed a more nuanced strategy. For example, in April, 2016 (a month prior to a scheduled lecture by Yiannopolus), the university administration sent an email to all students. The email started by reminding students that they "have the right to physical safety," and that they "should *never* be the subject of physical abuse, threats of violence, harassment, or intimidation." The email went on to say:

> *"University policies and the law do not protect us from being offended or from feeling uncomfortable with opposing views. The University supports you in challenging speech or expression you find offensive by exercising your own free speech rights, but we encourage and promote civility and respect in every exchange.*

The University does not, and cannot, censor content, but we can facilitate alternative events, programs, or expressions proposed by other members of the UCSB community which are conducted within our regulations."

This response by the UCSB administration acknowledges that threats to safety will not be tolerated, but implies that being offended by communication is not the equivalent of violence. Rather than banning potentially offensive communication, the university offers to facilitate events that can ease tension or potentially create dialogue. This approach appears to be better aligned with the ideals of public deliberation. Though more research is needed, it is reasonable to think that this is approach is more likely to both embrace goals of diversity of opinion and those of educating students.

Classroom Instruction

Although university policy and communications likely have effects on students' willingness to express themselves, each classroom has its own environment that is created by the instructor. Arguably, classroom instructors have an even greater opportunity to facilitate discussions about difficult topics and teach students how to engage in productive debate. Furthermore, college is often when students learn how to form arguments, evaluate evidence, and engage in rational deliberation with those who hold alternative viewpoints. Research has found that when the norms of debate are clear, students are more willing to engage in interracial conversations (Avery, Richeson, Hebl, & Ambady, 2009). In class room discussions about PC related topics, instructors can not only make the norms of discussion very clear, but they can carefully facilitate discussions in such a way that encourages rational deliberation (Avery & Steingard, 2008). Study 3 indicates that liberals may think less complexly about issues to which PC is relevant, and this complexity may be further

diminished by PC primes. Though Study 3 was conducted within the context of a dyadic conversation with a peer, rather than in a classroom setting, one can imagine that the findings should also apply to communications in a classroom. If college instructors have a goal of helping their students think more complexly about societal issues, then they should consider how their own communication about these subjects may potentially prime simplistic or complex thinking.

Ultimately, each instructor needs to consider their goals as a teacher, the content of the course they are teaching, the demographic makeup of their students, and the larger socio-political climate in which their course takes place. For example, the results of Study 2 showed that when an instructor used PC language, conservative students were less likely to report that they would express their opinions. On a campus that is predominantly liberal (as is UCSB), conservative students can offer a fresh perspective. On the other hand, at a university that is predominately White and/or conservative, an instructor might think that it is worth the tradeoff of silencing members of conservative majority if it encourages students from other backgrounds to speak up. The findings of this dissertation are not prescriptive for classroom instructors. Instead, they should merely be used as one piece of information for instructors so that they are aware of the potential effects their communication might have on students; from there, it is their decision about how to weight that information.

Limitations

As with any study, the studies presented in this dissertation have limitations that must be considered. Though many of these limitations have been addressed in the individual studies, those limitations that are relevant to all three studies and are addressed below.

Interpretations of PC

One of the features of PC, as laid out in Chapter I and by Robinson and Reid (2016a), is what is and what is not consistent with liberal PC ideology is open to interpretation, and this allows for strategic attributions of offense and social sanctioning (Robinson & Reid, 2015). This subjectivity is used on the left (in making determinations about who has violated the prescriptions of PC) and on the right (in making accusations about who is "too PC"). These strategic interpretations of norm adherence and violation makes PC ideologies useful in intergroup competition. However, if PC is subjective, then that means that participants' evaluations of what is PC likely vary. Across all three studies, I assumed (perhaps falsely) that participants subscribed to the popular definitions of PC (i.e., avoiding offending marginalized groups; Lalonde, Doan, & Patterson, 2000). However, because "PC" is also often invoked as an insult used against the left (Lalonde et al., 2000; Lea, 2009; Wilson, 1995), using the term "PC" had the potential to prime participants' negative attitudes towards those who use PC as a pejorative term (c.f. Strauts & Blanton, 2015). For this reason, I chose to avoid using the term "PC" in questions and stimuli. However, it is possible that participants did not interpret the items in Study 1 regarding speakers to ban, the code words in the stimulus in Study 2, and the position presented by the confederate in Study 3, as PC. A manipulation check at the end of each study would have been useful for establishing the extent to which participants viewed the content of studies as "PC". In addition to asking participants how "PC" they find the stimuli to be, they should also be asked their perceptions of what it is about the content that makes it PC. This will allow for a more nuanced understanding into how students understand PC. Follow up studies will include such manipulation checks. Future research should examine what attitudes and language

participants' associate with PC, as well as evaluate participants' interpretation of experimental stimuli.

A second important issue to consider is the content of the PC norms represented in this study. As Robinson and Reid (2016a) argue, and as I have argued throughout this dissertation, PC occurs on both the right and left. However, this dissertation primarily tested reactions to liberal PC normative language and attitudes. This was partially done due to the ideological leanings of those in the sample (see next section on sampling). For example, the purpose of Study 2 was to test the extent to which PC ideologies are primed using normative language. Therefore, it would make sense to prime the dominant ideology within the population, and the university where data collection took place is predominantly liberal. Furthermore, issues of liberal PC are not just salient on UCSB's campus, but have been recognized as an issue on college campuses more generally. There has been a rise in media attention paid to PC at universities (e.g., Chait, 2015; Lukianoff & Haidt, 2015; Weigel, 2016), and these popular press pieces adopt the popular definition of PC (avoiding offending marginalized groups). By testing PC within the liberal context, the findings from this dissertation are more likely to make a practical contribution to the current debate about PC in academia (which is where much of the debate about PC had occurred historically: Berman, 1992; Hughes, 2010; Lukianoff & Haidt, 2015; Loury, 1994; Weigel, 2016). It is common for researchers to examine communication and psychological phenomena in the contexts that are most socially relevant. For example, even though all people hold stereotypes, and that includes both negative and positive stereotypes about socially powerful groups, much of the social psychological work about stereotyping has focused on negative stereotypes about minority groups and women (see Augoustinos & Walker, 1998). Similarly, though there are

extremist groups on the left, scholarship about extremism in the United States tends to be about right-wing supremacy groups (e.g., Gerstenfeld, Grant, & Chiang, 2003; McCann, 2010; Wojcieszak, 2010). That said, in order to test the claims made by Robinson and Reid (2016a) and in this dissertation, that the features of PC norms and their consequences should exist independent of content of the ideology, future research is needed.

Sample

Sampling issues call some of the findings from this study into question. Across all three studies, the sample was predominantly female. Women likely view PC issues differently than many men, as feminist movements have long been aligned with PC movements (Hughes, 2010; Robinson & Reid, 2015). In addition, though the sample in all three studies was ethnically diverse (and is similar to the racial makeup of the university as a whole; UCSB Campus Profile, 2017), the categorization procedures used in analyses in Studies 1 and 2 (White people, People of Color) may have generalized important variation among People of Color. The stimulus in Study 2 referred to a broad topic ("campus climate"), and the discussion participants had in Study 3 (banning speakers from campus) was also framed broadly, rather than specifying which speaker and exactly what position they would be discussing. Because of this, there was not a specific reason why People of Color should have responded differently based on their specific racial group. However, Blacks, Asian, Latinos, and other racial groups all have different political and social interests that are salient to their specific group (e.g., police brutality, immigration), so had these stimuli specified an issue that was relevant to one minority group but not another, then it would have been more important to measure differences between racial groups. In addition, I did not ask participants about their sexual orientation. This was relevant to Study 1, where

participants evaluated whether or not speakers who advocate certain positions should be allowed to come to campus. Of the three conservative positions, two of them dealt with issues relevant to sexual minorities (transgenderism as a mental illness and homosexual conversion therapy), so members of these groups may have felt particularly threatened by the thought of these speakers coming to campus.

In the future, research should examine how making group-specific issues salient affects the extent to which priming PC is predictive of opinion expression. PC serves a coalitional function of bonding minority groups and White liberals together (Ford, 2017; Haidt & Leo, 2016; Robinson & Reid, 2016a). However, this coalition is useful to groups to the extent that it serves their interests. In fact, research on competitive victimhood has found that minority groups often compete with each other for victimhood status in order to obtain resources and garner attention for their cause (Young & Sullivan, 2016). For example, there have been reports of Blacks supporting Trump's immigration restrictions as they blame Latino immigrants for unemployment problems (Nichols, 2017), and there has been a growing schism between some feminists and trans-acceptance advocates, as some feminists believe that gender is a social construction rather than biological, so it is not possible that a "woman" could be born in a man's body (Soh, 2017). In order to examine the extent to which PC serves this coalitional building function, researchers should examine how the effects of PC primes differ across issues types and across groups.

Most importantly in terms of sampling, the participants in these studies were largely liberal. Though a university that has a liberal student population and embraces the prescriptions of PC provides an interesting opportunity to examine the perceptions and behaviors of a group of people who generally endorse PC norms, the results may have

differed on a campus with more political diversity. Though there were participants who were more conservative relative to others in the sample, there were very few far-right conservatives. Most participants were either highly committed liberals or left-leaning moderates. This is not necessarily problematic for serving the purpose of this dissertation, as scholars have pointed out that those with extreme views are less concerned with the political ostracism associated with dissent, and political moderates may be the most likely to be silenced by PC, as they are still concerned about the acceptance of the community (Loury, 1994). Furthermore, because of the small proportion of conservatives at UCSB, it may be that those who are conservative are more prone to self-censorship than conservatives at more politically diverse universities. It is possible that at a more politically heterogeneous university, conservatives, who tend to adamantly reject PC norms, would not interpret PC language as a cue to hide their opinions, but instead would see it as an opportunity to engage in intergroup competition. Though this political sample was useful in testing the hypotheses guiding this dissertation, it is important to consider how normative attitudes and behaviors differ at universities that are more politically heterogeneous (i.e, less norm consensus), or homogenously conservative. In these more conservative environments, priming PC may not be related to the self-censorship of dissent, as students may be more accustomed to hearing dissent and debate.

Measuring Diversity of Opinion

An underlying assumption of this dissertation is that censoring groups (or causing them to self-censor) necessarily decreases diversity of opinion. In fact, one of the themes that emerged from texts created in Study 3 was about the extent to which universities are supposed to be a place where people engage with those with differing opinions. Anecdotally,

from reading the emails from Study 3, it appears that participants had different ideas of what "diversity of opinion" means. Some think it means hearing from people from different backgrounds (e.g., based on race or culture), and others think it means hearing from people with different political opinions. In the case of the former, this is a common argument in favor of PC: by silencing some members of dominant groups (Whites, men), it will open space for members of groups that have been silenced in the past. However, the data from Study 2 do not support this argument. The data show that PC did not encourage those from minority racial groups to speak up more than did non-PC language. However, the results of this dissertation rely on self-report measures (Studies 1 and 2) and a dyadic interaction (Study 3). It is possible that in classroom environment, where many people are sharing their views, Students of Color might choose not to express themselves if confronted with too many viewpoints from White people (Rotham, Lispet, & Nevitte, 2003; White, 2011). It is possible that particularly in environments where PC is non-normative or classrooms that are highly skewed in their demographic makeup (mostly White with only a few Students of Color; mostly men with only a few women), these PC primes may increase diversity of opinion. However, considering that the makeup of UCSB is racially diverse (UCSB Campus Profile, 2017), and politically fairly homogenous, these primes are unlikely to increase diversity of opinion in the classroom.

Sanctions

In the case of PC, violators often incur harsh and public social sanctions, and it is thought that fear of these sanctions prevents people from expressing positions inconsistent with PC ideology (Chait, 2015; Ford, 2017; Robinson & Reid, 2016a). However, the procedures of these three studies did not test the effects of fear of social sanctions. In Study

2, participants were asked to imagine that they were in a classroom setting. In Study 3, participants communicated with a person to whom they would remain anonymous. Even if they did disagree with their partner, there was no risk of receiving any kind of social sanction outside of the lab. These inductions did not test the argument that fear of sanctions is what causes self-censorship. A replication that asks participants about the extent to which they think they would be sanctioned for expressing their views would be a simple test of this prediction. However, the manipulations in these studies do show that a relatively weak induction (i.e., one that does not provide the possibility of severe sanctions) was still enough to produce an effect of PC primes on self-censorship (in Study 2) and decreased integrative complexity (in Study 3). This provides support for the argument that PC has broad authority.

A Model of PC in Intergroup Relations

This dissertation tested the how perceptions of group victimhood, support for censorship, opinion expression/self-censorship, and public deliberation are related to political partisanship, PC normative attitudes and behavior, the consequences these relationships have for political deliberation. Furthermore, these processes are embedded in a larger intergroup context that involves identity politics--in essence, conservatives and liberals have become opposed coalitions, and this coalitional process affects political deliberation. Combining the findings of this dissertation, along with the argument put forward by Robinson and Reid (2016a), I propose a model to be tested in future research on PC. See Figure 7. It proposes, as would other intergroup relations theorists, that there is a direct path from intergroup conflict to decreased public deliberation (i.e., more contentious and polarized conflicts should result in more simplistic attributions of right and wrong, less complex debate, and less rational deliberation). However, considering PC as a tool used to engage in intergroup

competition (Robinson & Reid, 2016a), I propose that there is also an indirect path through which intergroup conflict can still end up negatively affecting public deliberation. This model highlights the importance of PC in exacerbating intergroup conflict.

Direct Path: Intergroup conflict decreases public deliberation

If public deliberation is characterized by interactants acknowledging alternative viewpoints and being open to changing their mind (Carpini et al., 2005; Jackman & Sniderman, 2006), then one could expect that the more severe the intergroup conflict, the less deliberation those competing groups will engage in. Though democracy requires that groups disagree and argue for their positions, to the extent that parties are increasingly polarized (e.g., increasing polarization among political parties in the United States), group conflict becomes more intractable (e.g., Israeli-Palestinian conflict), or resorts to violent war (e.g., Hutus and Tutsis in Rwanda), and these disagreements are unlikely to be expressed in a way that are conducive to the democratic process. For example, research on competitive victimhood has found that groups in conflict that continue to blame each other for their plight are less likely to recognize the legitimacy of their opponents' arguments (Noor et al., 2012), and ultimately less likely to engage in the reconciliation process (Young & Sullivan, 2016). This is an indication that public deliberation, which requires acknowledging the others' perspective and integrating viewpoints in order to create solutions to problems, is unlikely in severe intergroup conflict. Essentially, the worse the political conflict, the less public deliberation rivaling groups will engage in. However, as I propose below, the direct path from intergroup conflict to decreased public deliberation is not the only route.

Indirect Path

The contribution that PC makes to our understanding intergroup relations beyond other theories (e.g., SIT, SCT) is by highlighting an important indirect route that intergroup conflict can have on public deliberation. Each of these paths is explicated below, but broadly speaking, based on the arguments of Robinson and Reid (2016a) and the findings of this dissertation, I propose that increased group conflict/polarization leads to a rise in PC ideologies (characterized by the features laid out in Chapter I). These ideologies are enforced on both ingroup and outgroup members, and are advocated in such way that those who follow the ideology's prescriptions are considered moral, and those who deviate from its prescriptions are considered immoral. This can lead to even greater polarization as it causes political moderates to either choose sides or stay silent. This means that PC ideologies have a great deal of authority. This ideological authority affects opinion expression such that those who hold discordant views express their opinion less and those who agree express their opinions more (this is moderated by ideological primes, ideological commitment, and individual differences). Last, opinion expression affects public deliberation (also moderated by ideological primes, ideological commitment, and individual differences).

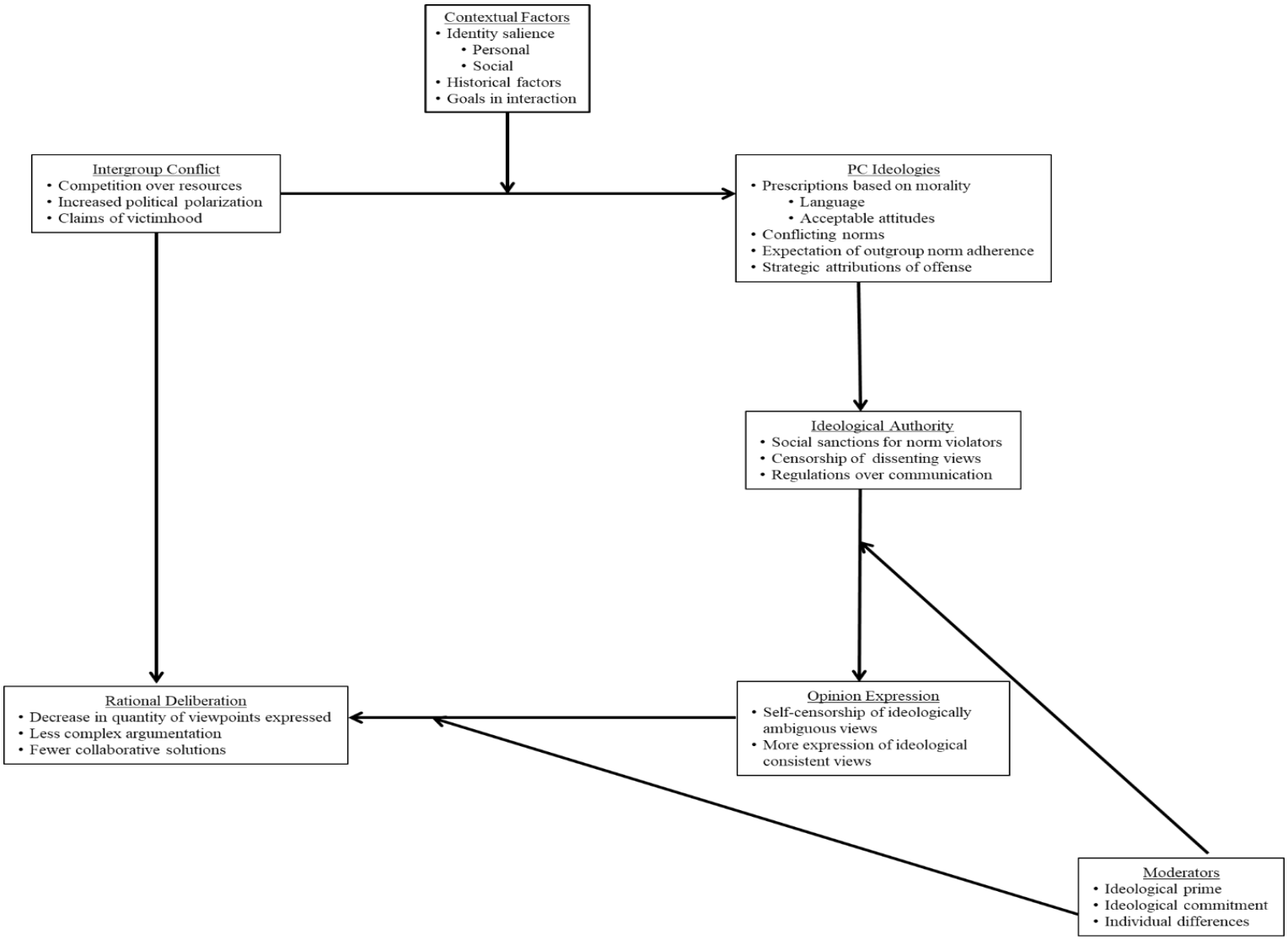

Figure 6. Proposed model, predicting indirect effects of intergroup conflict on public deliberation, via PC

Contextual factors. In line with the arguments of Robinson and Reid (2016a), I argue below that PC ideologies are more likely to come to rise during times of greater intergroup conflict. However, there are contextual factors that may moderate this relationship, particularly when examining singular interactions. For example, though social, rather than personal. identities are more likely to be salient during times of increased intergroup conflict (Tajfel, 1981; Tajfel & Turner, 1979), it may be possible that in an interpersonal interaction, personal identities can be made salient. Another factor that likely moderates the relationship is the history between the groups. For example, groups that have a history of violence may be more likely to deal with disputes with physical force, rather than engaging in conflict through the control of discourse, and groups may also differ in their history of power dynamics (have the two groups historically had unequal power or more equal power?). Last, in any given interaction, participants have different goals. These goals may be in service of engaging in intergroup competition, but there may be other times when the goal of an interaction may be interpersonal in nature (e.g., relationship building), or where interactants desire to improve intergroup relations (e.g., intergroup dialogue groups, classroom discussions, reconciliation processes). These contextual factors are likely influential in moderating the extent to which intergroup conflict leads to PC ideologies and the salience of such ideologies.

PC ideologies. According to Tajfel (1982) stereotypes serve an ideological function of explaining the current status of the group, as well as to justify actions against outgroups. As I argued in Chapter I, PC ideologies (meaning those ideologies that include the features of PC: prescriptions about language, conflicting norms, strategic attributions of offense,

social influence to adhere to prescriptions, censorship) rely on these functions of stereotypes, but in addition, also rely on attributions of morality/immorality. Adherents to PC, on both the right and left, feel threatened by the outgroup, and use this threat to justify actions against rivals. Robinson and Reid (2016a) argue that because PC serves this ideological function, it should increase during times of increased group conflict. Because PC is about control of political power through controlling discourse, it should increase during times of greater political polarization. However, though this process can occur between any groups, with any kind of threat, and with any actions taken against the outgroup, PC is more specific, in particular, the action taken against the outgroup is in the form of control of discourse (censorship, encouraging self-censorship).

Ideological authority. Because PC ideologies are grounded in a sense of morality, it allows groups to justify expecting both ingroup and outgroup members alike to adhere to PC. Any violation of PC, whether small or large, intentional or unintentional, can be met with accusations of immorality (Chait, 2015; Robinson & Reid, 2016a). Furthermore, because many of the normative prescriptions of PC ideologies conflict with each other or are in some other way open to interpretation, PC violations can be strategically sanctioned (Robinson & Reid, 2015). In line with Turner's (2005) three process model of persuasion, for those who do not subscribe to PC ideology, they may act in line with its prescriptions because an authority requires it (e.g., classroom policies regarding language use) or through coercion (e.g., fear of social sanctioning for non-adherence). This gives PC ideologies a great amount authority, above and beyond other group level prescriptions, as they are enforced on both ingroup members and outgroup members. This strategic enforcement not only means that people need to avoid expressing ideas that directly oppose PC ideology (e.g., using racial

slurs) but also those that may be interpreted as opposing PC. In this way, opinion expression is likely to decrease among those who hold partially discordant views.

Opinion expression. Because of PC's ideological authority, for those who endorse the ideology, they should be more likely to express their opinions about PC relevant issues. The opinion expression of those who hold ideologically inconsistent views, but are unconcerned about social sanctions (i.e., Noelle-Neumann's (1974) "hard core" individuals), should not be affected by PC, and should be likely to express their views independent of opinion climate (see Moderators). However, for those who hold views inconsistent with PC or who are uncertain about how their positions will be interpreted, but are also concerned about social sanctions, maintaining their identity as moral individuals, or want to remain a member of the community, they should be less likely to express their views.

Public deliberation. Opinion expression or inhibition is an important factor in public deliberation, in terms of the number of ideas that are expressed, to what extent the complexity of social issues is considered, and how likely groups are to come up with collaborative solutions to societal problems. When disagreements are attributed to moral failings rather than legitimate differences of opinion, then debate necessarily becomes less complete and ultimately leads to poorer decision making (Chambers, 2003). If public deliberation is the key component to a working democracy, then any ideological perspective or set of norms that inhibit this deliberation is cause for concern. To the extent that PC acts as a dogma that gives adherents simple explanations for right and wrong, then those who subscribe to these ideologies (on both the right and left) are less likely to engage in true public deliberation.

Moderators. Last, the path between ideological authority and opinion expression, and the path between opinion expression and public deliberation should be moderated by three factors. First, ideological primes give listeners important insight into what is the dominant ideology in that particular environment. As was found in Study 2, ideological primes should increase the effect of PC on opinion expression, and as was found in Study 3, should increase the effect of PC on integrative complexity (i.e., public deliberation). Second, these relationships should be moderated by ideological commitment. For those who strongly oppose PC (i.e., are committed to another ideology), they should not engage in self-censorship, even when in a PC dominated environment. Instead, it will be political moderates who are the most-likely to self-censor, as PC ideologies' simplistic prescriptions of right and wrong do not leave room for positions in between. Last, there are individual differences in the extent to which people are likely to self-censor (WTSC: Hayes et al., 2005) or the extent to which they are concerned with social sanctions (FSI: Hayes et al., 2011), and this too affects opinion expression and self-censorship.

Conclusion

The studies presented in this dissertation are a first step in broadening our understanding of the connection between PC, endorsement of censorship, self-censorship, and public deliberation. This dissertation has made a contribution, both theoretically, and practically. Theoretically, these findings have highlighted the importance of understanding the role of control of discourse, free expression, and self-censorship in intergroup relations. From a practical perspective, after 30 years of scholars, journalists, and politicians debating PC's advantages and disadvantages on college campuses, the results of this dissertation indicate that identity politics play an important role in students' support for censorship, that

PC ideological primes are related to self-censorship, and that more liberal students may have difficulty discussing PC relevant topics in complex ways. Taken together, these findings provide empirical support for the decades long argument that PC can create a chilling effect on political deliberation.

References

Abe, J. A. (2011).Changes in Alan Greenspan's language use across the economic cycle: A text analysis of his testimonies. *Journal of Language and Social Psychology, 30*, 212-223. doi: 10.1177/0261927X10397152

Augoustinos, M., & Walker, I. (1998). The construction of stereotypes within social psychology: From social cognition to ideology. *Theory and Psychology, 8,* 629-652. doi: 10.1177/0959354398085003

Al-Gharbi, M., & Haidt, J. (2017, July 8). It's disadvantaged group the suffer most when free speech is curtailed on campus. *The Atlantic.* Retrieved from: https://www.theatlantic.com/politics/archive/2017/07/free-speech-campus/532965/

Andris, C., Less, D., Hamilton, M., Martino, M., Gunning, C., & Selden, J. (2015). The rise of partisanship and super-cooperators in the U.S. House of Representatives. *Public Library of Science, One, 10*.

Ardener, E. (1978). Some outstanding problems in the analysis of events. In G. Schwinner (Ed.), *The Yearbook of Symbolic Anthropology* (pp. 103-121). London: Hurst.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs, 70,* 1-70. doi: 10.1037/h0093718

Avery, D. R., & Steingard, D. S. (2008). Achieving political trans-correctness: Integrating sensitivity and authenticity in diversity and management education. *Journal of Management Education, 32,* 269-293. doi: 10.1177/1052562907305557

Bacon, P. (2017, August 14). Charlottesville and the rise of White identity politics. *FiveThirtyEight.* Retrieved from: https://fivethirtyeight.com/features/charlottesville-and-the-rise-of-white-identity-politics/

Bakshy, E., Messing, S., & Adamic, L. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science, 348,* 1130-1132.

Berman. (1992). Introduction: The debate and its origins. In. P. Berman (Ed.), *Debating P.C.* (pp. 1-29) New York, NY: Dell Publishing.

Boyd, R. L. (2017a). MEH: Meaning Extraction Helper (Version 1.4.15) [Software]. Available from http://meh.ryanb.cc

Boyd, R. L. (2017b). Psychological text analysis in the digital humanities. *Data Analytics in Digital Humanities,* 161-189. Springer International Publishing.

Boyd, R. L., & Pennebaker, J.W. (2015). A way with words: using language for psychological science in the modern era. In: Dimofte, C., Haugtvedt, C., Yalch, R. (Eds.), *Consumer psychology in a social media world*. (pp. 222–236). New York, NY: Routledge Publishers.

Boyd, R. L., Wilson, R. S., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015). Value in words: using language to evaluated and understand personal values. *Proceedings of the Ninth International AAAI Conference on Web and Social Media,* 31-40.

Branscombe, N. R. (1998). Thinking about one's gender group's privileges or disadvantages: Consequences for well-being in women and men. *British Journal of Social Psychology, 37,* 167-184. doi: 10.1111/j.2044-8309.1998.tb01163.x

Brown, T. (2006). *Confirmatory Factor Analysis for Applied Research*. Guilford Press, New York, NY: Guilford press.

Brundidge, J., Reid, S. A., Choi, S., & Muddiman, A. (2014). The "deliberative digital divide:" Opinion leadership and integrative complexity in the U.S. political

blogosphere. *Political Psychology, 35,* 741-755. doi: 10.1111/pops.12201

Carnaghi, A. & Maass, A. (2007). In-group and out-group perspectives in the use of derogatory group labels: Gay versus fag. *Journal of Language and Social Psychology, 26,* 142-156. doi: 10.1177/0261927X07300077

Carpini, M. X. D., Cook, F. L., & Jacobs, L. R. (200). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science, 7,* 315-344. doi: 10.1146/annurev.polisci.7.121003.091630

Chait, J. (2015, January 27). Not a very P.C. thing to say. *New York Magazine*. Retrieved from: http://nymag.com/daily/intelligencer/2015/01/not-a-very-pc-thing-to-say.html

Chambers, S. (2003). Deliberation democratic theory. *Annual Review of Political Science, 6,* 307-326. doi: 10.1146/annurev.polisci.6.121901.085538

Chung, C. K., & Pennebaker, J. W. (2008). Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality, 42,* 96-132. doi: 10.1016/j.jrp.2007.04.006

Chung, C. K., & Pennebaker, J. W. (2014). Using computerized text analysis to track social processes. In T. M. Holtgraves (Ed.), *The Oxford Handbook of Language and Social Psychology.* (pp. 219-232). New York, NY: Oxford University Press.

Collins, K. A., & Clément, R. (2012). Language and prejudice: Direct and moderated effects. *Journal of Language and Social Psychology, 31,* 376-396. doi: 10.1177/0261927X12446611

Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. (2014). Automated integrative complexity. *Political Psychological, 35*, 603-624. doi: 10.1111/pops.12021

Conway, L. G., Gornick, L. J., Houck, S. C., Anderson, C., Stockert, J., Sessoms, D., & McCue, K. (2016). Are conservatives really more simple-minded than liberals? The domain specificity of complex thinking. *Political Psychology, 37*, 777-798. doi: 10.1111/pops.12304

Dahlberg, L. (2005). The Habermasian public sphere: Taking difference seriously? *Theory and Society, 34*, 111-136. doi: 10.1007/s11186-005-0155-z

Danescu-Neculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of power: Language effects power and differences in social interaction. In *Proceedings of the 21st International World Wide Web Conference*. New York, NY: ACM. Retrieved from: http://arxiv.org/pdf/1112.3670.pdf

Delton, J. (2017, August 22). When 'free speech' becomes a political weapon. *The Washington Post.* Retreived from: https://www.washingtonpost.com/news/made-by-history/wp/2017/08/22/when-free-speech-becomes-a-political-weapon/?utm_term=.cf95ce4f86eb

Demby, G. (2013, January 30). In immigration debate, 'undocumented' vs. 'illegal' is more than just semantics. *NPR*. Retrieved from: http://www.npr.org/sections/itsallpolitics/2013/01/30/170677880/in-immigration-debate-undocumented-vs-illegal-is-more-than-just-semantics

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In B. H. Ross (Series Ed.) & D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin

(Eds.), *Psychology of learning and motivation, Vol. 50*: Moral judgment and decision making (pp. 307-338). San Diego, CA: Academic Press.

Dragojevic, D., Gasiorek, J., & Giles, H. (2016). Accommodative strategies as core of the theory. In H. Giles (Ed.), *Communication accommodation theory: Negotiating personal relationships and social identities across contexts.* (pp. 36-59). Cambridge, United Kingdom: Cambridge University Press.

Dragojevic, M., & Giles, H. (2014). Language and interpersonal communication: Their intergroup dynamics. In C. R. Berger (Ed.), *Interpersonal communication* (pp. 29-51). Berlin, Germany: De Gruyter.

Dreger, A. (2015). *Galileo's middle finger: Heretics, activists, and the search for justice*. New York, NY: Penguin Press.

D'Souza, D. (1992). The big chill? Interview with Dinesh D'Souza. In. P. Berman (Ed.), *Debating P.C.*(pp. 29-39). New York, NY: Dell Publishing

Dunn, D. S., & Andrews, E. E. (2015). Person-first and identity-first language: Developing psychologists' cultural competence using disability language. *American Psychologist, 70*, 255-264. doi: 10.1037/a0038636

Ford, B. R. (2017). Political correctness. In H. Giles & J. Harwood (Eds.) *Oxford Encyclopedia of Intergroup Communication*. (pp TBD). New York, NY: Oxford University Press.

Gabriel, U., & Gyax, P. (2016a). Gender and linguistic sexism. In H. Giles & A. Maass (Eds.) *Advances in Intergroup Communication*. (pp. 177-192). New York, NY: Peter Lang Publishing.

Gallois, C. & Pittman, J. (1995). Social psychological approaches to using natural language texts. *Journal of Language and Social Psychology*. *33,* 928-940. DOI: 10.1177/0261927X95141001

Gasiorek, J. (2016a). Theoretical perspectives on interpersonal adjustments in language and communication. In H. Giles (Ed.), *Communication accommodation theory: Negotiating personal relationships and social identities across contexts.* (pp. 13-35). Cambridge, United Kingdom: Cambridge University Press.

Gasiorek, J. (2016b). The "dark side" of CAT: Nonaccommodation. In H. Giles (Ed.), *Communication accommodation theory: Negotiating personal relationships and social identities across contexts.* (pp. 85-104). Cambridge, United Kingdom: Cambridge University Press.

Gerstenfeld, P. B., Grant, D. R., Chiang, C. P. (2003). Hate online: A content analysis of extremist internet sites. *Analyses of Social Issues and Public Policy*, *3*, 29-44. doi: 10.1111/j.1530-2415.2003.00013.x

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, and N. Coupland (eds.), *Contexts of accommodation: Developments in applied linguistics* (pp. 1–68). New York, NY: Cambridge, University Press.

Glynn, C. J., Hayes A. F., & Shanahan, J. (1997). Perceived support for one's opinions and willingness to speak out: A meta-analysis of survey studies on the spiral of silence. *The Public Opinion Quarterly, 61*, 452-463. doi: 10.1086/297808

Golec De Zavala, A., Cisak, A., & Wesolowska, E. (2010). Political conservatism, need for cognition, and intergroup hostility. *Political Psychology*, *31,* 521–541. doi:

10.1111/j.1467-9221.2010.00767.x

Graesser, A. C., McNamara, D. S., Louwerse, M.M. et al. (2004). Coh-metrixs: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*, 193-202. doi:10.3758/BF03195564

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberal and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*, 1029-1046. doi: 10.1037/a0015141

Grimmer, J., & Sterwart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*, 267-297. doi: 10.1093/pan/mps028

Habermas, J. (1984). *Theory of communicative action* (2 vols.). Boston: Beacon Press.

Habermas, J. (1989). *The structural transformation of the public sphere*. Cambridge, MA: MIT.

Haertl, B. E., & McCarthy, P. M. (2010). Differential linguistic features in U.S. immigration newspaper articles: A constrastive corpus analysis using the Gramulator. *Artificial Intelligence, 24*, 349-350.

Haidt, J. (2008). Morality. *Perspectives in Psychological Sicence, 3*, 65-72. doi: 10.1111/j.1745-6916.2008.00063.x

Haidt, J. (2012). *The Righteous Mind: Why good people are divided by politics and religion*. New York, NY: Pantheon Books.

Haidt, J. & Leo, J. (2016, February 3). A conversation with Jonathon Haidt. Retrieved from: http://www.mindingthecampus.org/2016/02/a-conversation-with-jonathan-haidt

Haidt, J., Rosenberg, E., & Hom, H. (2003). Differentiating diversities: Moral diversity is

not like other kinds. *Journal of Applied Social Psychology, 33,* 1-36. doi: 10.1111/j.1559-1816.2003.tb02071.x

Hayes, A. F. (2007). Exploring forms of self-censorship: On the spiral of silence and the use of opinion expression avoidance strategies. *Journal of Communication, 57,* 785-802. doi: 10.1111/j.1460-2466.2007.00368.x

Hayes, A. F. (2013*). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press

Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Validating the Willingness to Self-Censor scale: Individual differences in the effect of the climate of opinion on expression. *International Journal of Public Opinion Research, 17,* 443-455. doi: 10.1093/ijpor/edh072

Hayes, A. F., Matthes, J., & Eveland, W. P. (2013). Stimulating the quasi-statistical organ: Fear of social isolation motivates the quest for knowledge of the opinion climate. *Communication Research, 40,* 1-24. doi: 10.1177/0093650211428608

Hayes, A. F., Scheufele, D., & Huge, M. (2006). Nonparticipation as self-censorship: Publicly observable political activity in a polarized opinion climate. *Political Behavior, 28,* 259-283. doi: 10.1007/s11109-006-9008-3

Hentoff (1992). "Speech codes" on the campus and problems of free speech. In. P. Berman (Ed.), *Debating P.C.*(pp. 215-224). New York, NY: Dell Publishing

Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication Theory, 16,* 7-30. doi: 10.1111/j.1468-2885.2006.00003.x

Hofstadter, R. (1965). *The paranoid style in American politics and other essays*. New York,

NY: Alfred A Knopf.

Hornsey, M. J., & Imani, A. (2004). Criticizing groups from the inside and the outside: An identity perspective on the intergroup sensitivity effect. *Personality and Social Psychology Bulletin, 30,* 365-383. doi: 10.1177/0146167203261295

Hornsey, M. J., Trembath, M., & Gunthorpe, S. (2004). 'You can criticize because you care': Identity attachment, constructiveness, and the intergroup sensitivity effect. *European Journal of Social Psychology, 34*, 499-518. doi: 10.1002/ejsp.212

Hughes, G. (2010). *Political Correctness: A history of semantics and culture*. Chichester, UK: Wiley & Sons.

Hurwtiz, J. & Peffley, M. (2005). Playing the race card in the post-Willie Horton era: The impact of racialized code words on support for punitive crime policy. *Public Opinion Quarterly, 69,* 99-112. doi: 10.1093/poq/nfi004

Ireland, M. E., & Mehl, M. R. (2014). Natural language use as a marker of personality. In T.M. Holtgraves (Ed.), *The Oxford Handbook of Language and Social Psychology.* (pp. 201-218). New York, NY: Oxford University Press.

Ireland, M. E., & Penebaker, J. W. (2010). Language style matching in reading and writing. *Journal of Personality and Social Psychology, 99,* 549-571. DOI: 10.1037/a0020386

Iyer, A., Jetten, J., Branscobme, N. R., Jackson, S., & Youngberg, C. (2014). The difficulty of recognizing less obvious forms of group-based discrimination. *Group Processes and Intergroup Relations, 17*, 577-589. doi: 10.1177/1368430214522139

Jackman, S., & Sniderman, P. M. (2006). The limits of deliberative discussion: A model of everyday political arguments. *The Journal of Politics, 68*, 272-283. doi: 10.1111/j.1468-2508.2006.00405.x

Jetten, J., Schmitt, M. T., & Branscobme, N. R. (2012). Rebels *without* a cause: Discrimination appraised as legitimate harms group commitment. *Group Processes and Intergroup Relations, 16,* 159-172. doi: 10.1177/1368430212445075

Jones, R. P. (2017). Republican more likely to saw White Americans—rather than Black Americans—face discrimination. Public Religion Research Institute. Retrieved from: https://www.prri.org/spotlight/republicans-white-black-reverse-discrimination/

Jost, J. T., Glaser, J., Kruglanski, A. W., Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129,* 339-375. doi: 10.1037/0033-2909.129.3.339

Kipnis, L. (2015, May 29). My title IX inquisition. *The Chronical of Higher Education*. Retrieved from: http://chronicle.com/article/My-Title-IX-Inquisition/230489?key=S25xdAJqZHweM35rZm0VNzsAPSY8Mkl5MHpPaipybltcGQ%3D%3D

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480-498. doi: 10.1037/0033-2909.108.3.480

Lambe, J. L. (2008). The structure of censorship attitudes. *Communication Law and Policy, 13*, 485-506. doi: 10.1080/10811680802388923

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *The American Political Science Review, 97,* 311-331. doi: 10.1017/S0003055403000698

Lea, J. (2009). *Political correctness and higher education: British and American perspectives*. New York, NY: Routledge.

Leets, L., & Giles, H. (1997). Words as weapons—when do they wound? Investigations of

harmful speech. *Human Communication Research, 24,* 260-301. doi: 10.1111/j.1468-2958.1997.tb00415.x

LeFebvre, L., LeFebvre, L., Blackburn, K., Boyd, R. L. (2015). Student estimates of public speaking competency: The Meaning Extraction Helper and video self-evaluation. *Communication Education, 64,* 261-279. doi: 10.1080/03634523.2015.1014384

Loury, G. C. (1994). Self-censorship in public discourse: A theory of "political correctness" and related phenomena. *Rationality and Society, 6,* 428-461. doi: 10.1177/1043463194006004002

Lukianoff, G. & Haidt, J. (2015, September). The coddling of the American mind. *The Atlantic.* Retrieved from: http://www.theatlantic.com/magazine/archive/2015/09/the-coddling-of-the-american-mind/399356/

Maass, A., Suitner, C., & Merkel, E. (2014). Does political correctness make (social) sense? In. J. Forgas, O. Vincze, & J. László (Eds.), *Social cognition and communication* (pp. 331-346). New York, NY: Psychology Press.

Marques, J. F. (2009). How politically correct is political correctness?: A SWOT analysis of this phenomenon. *Business & Society*, *48*. 257-266. doi: 10.1177/0007650307307155

Martin, R., Martin, P. Y., Smith, J. R., & Hewstone, M. (2007). Majority versus minority influence and prediction of behavior intentions and behavior. *Journal of Experimental Social Psychology, 43,* 763-771. doi: 10.1016/j.jesp.2006.06.006

Matthes, J., Hayes, A. F., Rojas, H., Shen, F., Min, S. J., & Dylko, I. B. (2012). Exemplifying a dispositional approach to cross-cultural spiral of silence research: Fear of social isolation and the inclination to self-censor. *International Journal of Public Opinion Research*, *24,* 287-305. doi; 10.1093/ijpor/eds015

McCann, S. J. (2010). Authoritarianism, conservativism, racial diversity threat, and the state of distribution of hate groups. *The Journal of Psychology, 144*, 37-60. doi: 10.1080/00223980903356065

Mendelberg, T. (2002). The deliberative citizen: Theory and evidence. In M. D. X. Carpini, L. Huddy, & R. Shapiro (Eds.) *Research in micropolitics: Political decisionmaking, deliberation, and participation.* (pp 151-93). Greenwich, CT: JAI Press

Mendelberg, T., & Oleske, J. (2000). *Political Communication, 17,* 161-191. doi: 10.1080/105846000198468

Miller, D. T., & Morrison, K. R. (2008). Expressing deviant opinions: Believing you are in the majority helps. *Journal of Experimental Social Psychology, 45,* 740-747. doi: 10.1016/j.jesp.2009.04.008

Monroe, B. L., Colaresi, M.P., & Quinn, K. M. (2008). Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis, 16,* 372-403. doi: 10.1093/pan/mpn018

Morphew, C. C., & Hartley, M. (2006). Mission statements: A thematic analysis of rhetoric across institutional type. *The Journal of Higher Education, 77,* 456-471. doi: 10.1080/00221546.2006.11778934

Morris, S. (2001). Political correctness. *Journal of Political Economy, 109,* 231-265. doi: 10.1086/319554

Moscovici, S. (1980). Toward a theory of conversion behavior. *Advances in Experimental Social Psychology, 13,* 209-239. doi: 10.1016/S0065-2601(08)60133-1

Moscovici, S., & Personnaz, B. (1980). Studies in social influence: minority influence and conversion behavior in a perceptual task. *Journal of Experimental Social Psychology, 16,* 270-282. doi: 10.1016/0022-1031(80)90070-0

Mulac, A., Bradac, J. J., & Gibbons, P. (2001). Empirical support for the gender-as-culture hypothesis: An intercultural analysis of male/female language differences. *Human Communication Research, 27,* 121-152. doi: 10.1111/j.1468-2958.2001.tb00778.x

Mulac, A., Studley, L. B., Wienmann, J. M., & Bradac, J. J. (1987). Male/female gaze in same-sex and mixed-sex dyads gender-linked differences in mutual influence. *Human Communication Research, 13*, 323-343. doi: 10.1111/j.1468-2958.1987.tb00108.x

Mullen, B., Atkins, J.L., Champion, D. S., Edwards, C., Hardy, D., Story, J.E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, *21*, 262-283. doi: 10.1016/0022-1031(85)90020-4

Mullen, B., & Smyth, J.M. (2004). Immigrant suicide rates as a function of ethnophaulisms: Hate speech predicts death. *Psychosomatic Medicine, 66,* 343-348.

Namy, L. L., Nygaard, L. C., Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology, 21*, 422-432. doi: 10.1177/026192702237958

Nichols, J. (2017, January 15). Immigration should not divide Black and Latinx people. *Al-Jazeera*. Retrieved from: http://www.aljazeera.com/indepth/opinion/2017/01/immigration-divide-black-latinx-people-170112150841898.html

Noelle-Neumann, E. (1974). The Spiral of Silence: A theory of public opinion. *Journal of*

Communication, 24, 43–51. doi: 10.1111/j.1460-2466.1974.tb00367.x

Noelle-Neumann, E. (1991). The theory of public opinion: The concept of the spiral of silence. *Annals of the International Communication Association, 14*, 256-287. doi: 10.1080/23808985.1991.11678790

Noor, M., Shnabel, N., Samer, H., & Nadler, A. (2012). When suffering begets suffering: The psychology of competitive victimhood between adversarial groups in violent conflict. *Personality and Social Psychology Review, 16*, 351-374. doi: 10.1177/1088868312440048

Orbe, M. P. (1998). From the standpoint(s) of traditionally muted groups: Explicating a co-cultural communication theoretical model. *Communication Theory, 8*, 1-26. doi: 10.1111/j.1468-2885.1998.tb00209.x

Orbell, J. M., Van de Kragt, A. J., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*, 811-819. doi: 10.1037/0022-3514.54.5.811

Park, M. & Lah, K. (2017, February 2). Berkeley protests of Yiannopoulos caused $100,00 in damage. Retrieved from: http://www.cnn.com/2017/02/01/us/milo-yiannopoulos-berkeley/

Pennebaker, J. W. & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science*. *10*, 90-93. DOI: 10.1111/1467-8721.00123

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin

Peters, J. W., & Fuller, T. (2017, April 26). Ann Coulter says she will pull out of speech at

Berkeley. *The New York Times*. Retrieved from: https://www.nytimes.com/2017/04/26/us/ann-coulter-berkeley-speech.html?_r=0

Phillips, L. T., & Lowery, B. S. (2015). The hard-knock life? Whites claim hardships in response to racial inequality. *Journal of Experimental Social Psychology, 61,* 12-18. doi: 10.1016/j.jesp.2015.06.008

Plant, E. A., & Devine, P. G. (2001). Responses to other-imposed pro-Black pressure: acceptance or backlash? *Journal of Experimental Social Psychology, 37,* 486-501. doi: 10.1006/jesp.2001.1478

Plant, E. A., Devine, P. G., & Brazy, P. C. (2003). The bogus pipeline and motivations to respond without prejudice: Revisiting the fading and faking of racial prejudice. *Group Processes & Intergroup Relations, 6,* 187-200. doi: 10.1177/1368430203006002004

Potter, J. & Wetherell, M. (1995). Natural order: Why social psychologists should study (a constructed version of) natural language, and why some have not done so. *Journal of Language and Social Psychology. 33*, 928-940. DOI: 10.1177/0261927X95141012

Robinson, B., & Reid, S. A. (2015). *Motivated offense: Group membership, status competition, offense, and the desire to punish*. Presented at the annual meeting of the International Communication Association in San Juan, Puerto Rico.

Robinson, B. & Reid, S.A. (2016a). Political correctness. In H. Giles & A. Maass (Eds.) *Advances in Intergroup Communication*. (pp. 85-100). New York, NY: Peter Lang Publishing.

Robinson, B. & Reid, S. A. (2016b). Partisan code words affect policy opinions and group demographic estimation. Paper presented in the Intergroup Division of the International Communication Association, Fukuoka, Japan.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13,* 279-301. doi: 10.1016/0022-1031(77)90049-X

Rotham, S., Lispet, S. M., & Nvitte, N. (2003). Does enrollment diversity improve university education? *International Journal of Public Opinion Research, 15,* 8-26. doi: 10.1093/ijpor/15.1.8

Schlosser, E. (2015, June 3). I'm a liberal professor, and my liberal students terrify me. *Vox.* Retrieved from: http://www.vox.com/2015/6/3/8706323/college-professor-afraid

Schmader, R., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115,* 336-356. doi: 10.1037/0033-295X.115.2.336

Slapin, J. B., & Proksch, S. O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science, 52,* 705-722. doi: 10.1111/j.1540-5907.2008.00338.x

Soh, D. W. (2017, February 10). Are gender feminists and transgender activists undermining science? *Los Angeles Times*. Retrieved from: http://www.latimes.com/opinion/op-ed/la-oe-soh-trans-feminism-anti-science-20170210-story.html

Soliz, J., & Giles, H. (2014). Relational and identity processes in communication: A contextual and meta-analytical review of Communication Accommodation Theory. In E. L. Cohen (Ed.), *Communication Yearbook 38* (pp.107-144). Thousand Oaks, CA:

Sage. doi: 10.1017/CBO9781316226537.003

Spender, D. (1980). *Man made language*. London, UK: Routledge and Kegan Paul.

Spitzberg, B. (1991). A model of intercultural communication competence. In L. Samovar & R. Porter (Eds.), *Intercultural communication: A reader* (7th ed, pp. 347-359). Belmont, CA: Wadsworth

Strauts, E. & Blanton, H. (2015). That's not funny: Instrument validation of the concern for political correctness scale. *Personality and Individual Differences, 80,* 32-40. doi: 10.1016/j.paid.2015.02.012

Streufert, S., Suedfeld, P., & Driver, M. J. (1965). Conceptual structure, information search, and information utilization. *Journal of Personality and Social Psychology*, *2*, 736–740. doi: 10.1037/h0022679

Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). Conceptual/integrative complexity. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 393–400). NewYork: Cambridge University Press.

Sullivan, D., Landau, M. J., Branscombe, N. R., & Rothschild, Z. K. (2012). Competitive victimhood as a response to accusation of ingroup harm doing. *Journal of Personality and Social Psychology, 102,* 778-795. doi: 10.1037/a0026573

Sutton, R. M. (2010). The creative power of language in social cognition and intergroup relations. In H. Giles, S. A. Reid, & J. Harwood (Eds.), *The dynamics of intergroup communication* (pp. 105-115). New York, NY: Peter Lang.

Tajfel, H. (1981). *Human group and social categories: Studies in social psychology*. Cambridge, England: Cambridge University Press.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology,*

33, 1-29. doi: 10.1146/annurev.ps.33.020182.000245

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.

Twerwilleger, B., McCarthy, P. M., & Lamkin, T. (2011). Bias in hard news articles from Fox News and MSNBC: An empirical assessment using the Gramulator. *Artificial Intelligence, 24*, 361-362.

Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social Psychology*, *45*, 118–126. doi: 10.1037/0022-3514.45.1.118

Tetlock, P. E. (1984). Cognitive style and political belief systems in the British House of Commons. *Journal of Personality and Social Psychology*, *46*, 365–375. doi: 10.1037/0022-3514.46.2.365

Tetlock, P. E. (1985). Integrative complexity of the American and Soviet foreign policy rhetoric: A time-series analysis. *Journal of Personality and Social Psychology, 49,* 1565-1685. doi: 10.1037/0022-3514.49.6.1565

Tetlock, P. E. (1986). A value pluralism model of ideological reasoning. *Journal of Personality and Social Psychology, 50*, 819-827. doi: 10.1037/0022-3514.50.4.819

Turner, J. C. (2005). Explaining the nature of power: A three-process theory. *European Journal of Social Psychology, 35*, 1-22. doi: 10.1002/ejsp.244

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Blackwell.

UCSB Campus Profile. (2017, January). 2016-2017 campus profile: University of California-Santa Barbara. Retrieved from: http://bap.ucsb.edu/institutional.research/campus.profiles/campus.profiles.2016.17.pdf

Van Boven, L. (2000). Pluralistic ignorance and political correctness: The case of affirmative action. *Political Psychology, 21*, 267-276. doi: 10.1111/0162-895X.00187

Vinkour, A., & Burnstein, E. (1978). Depolarization of attitudes in groups. *Journal of Personal and Social Psychology, 36*, 872-885. Doi: 10.1037.0022-3514.36.8.872

Weigel, M. (2016, November 30). Political correctness: How the right invented a phantom enemy. *The Guardian*. Retrieved from: https://www.theguardian.com/us-news/2016/nov/30/political-correctness-how-the-right-invented-phantom-enemy-donald-trump

White, J.W. (2011). Resistance to classroom participation: Minority students, academic discourse, cultural conflicts, and issues of representation in whole class discussions. *Journal of Language, Identity, and Education, 10*, 250-265. doi: 10.1080/15348458.2011.598128

Whitney, D. C., & Wartella, E. (1992). Media coverage of the "political correctness" debate. *Journal of Communication, 42*, 83-95. doi: 10.1111/j.1460-2466.1992.tb00780.x

Wilkström, P. (2016). No one is "pro-politically correct": Positive construals of political correctness in Twitter conversations. *Nordic Journal of English Studies, 15*, 159-170. Retrieved from: http://ojs.ub.gu.se/ojs/index.php/njes/article/view/3532/2962

Wilson, J. K. (1995). *The myth of political correctness: A conservative attack on higher education.* Durham, NC: Duke University Press.

Wojcieszak, M. (2010). “Don’t talk to me”: Effect of ideologically homogenous online groups and politically dissimilar offline ties on extremism. *New Media & Society, 12*, 637-655. doi: 10.1177/1461444809342775

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*, 913-934. doi: 10.1177/0013164413495237

Young, I. F., & Sullivan, D. (2016). Competitive victimhood: A review of the theoretical and empirical literature. *Current Opinion in Psychology, 11*, 30-34. doi: 10.1016/j.copsyc.2016.04.004

Appendix A

Note: Measures are listed in alphabetical order. Items within each measure are listed in the same order as was seen by participants. All questionnaires were administered using Qualtrics.

Encouragement of opinion expression (used in Study 2)

(scale anchors: 1 (not at all) to 7 (very much))

To what extent is the professor encouraging members of the following groups to share their views:

Political liberals

Political conservatives

Members of ethnic minorities

Members of sexual minorities

Endorsement of freedom of expression (used in Study 3)

(from the 2005 2005 & 2006 State of the First Amendment surveys)

Anti-government speech: (1 (strongly agree) to 5 (strongly disagree))

Newspapers should be allowed to feely criticize the U.S. military about its strategy and performance
Newspapers should honor government requests to withhold publishing information that might hurt efforts to win the war on terrorism
Even during wartime, the press should be allowed to public stories that criticize the actions of the government

Offensive speech: (1 (strongly agree) to 5 (strongly disagree))

Musicians should be allowed to sing songs with lyrics that others might find offensive
People should be allowed to say thing in public that might be offensive to racial groups

People should be allowed to say things in public that might be offensive to religious groups

Engagement in Conversation (used in Study 3)

(scale anchors: 1 (not at all) to 7 (very much))

Please respond to the following:

1. I was fully engaged in the conversation
2. I was interested in what my partner had to say
3. I enjoyed discussing this topic
4. The topic of the conversation is important to me
5. I feel strongly about this topic

Fear of Social Isolation (used in Studies 2 & 3)

(scale anchors: 1 (strongly disagree) to 5 (Strongly agree))

1. It is scary to think about not being invited to social gatherings by people I know
2. One of the worst things that could happen to me is to be excluded by people I know
3. It would bother me if no one wanted to be around me
4. I dislike feeling left out of social functions, parties, or other social gathering
5. It is important to me to fit into the group I am with
6. I worry about being isolated if people disagree with me
7. I don't worry about other people avoiding me (*reverse coded*)
8. I avoid telling other people what I think when I think there is a risk they'll avoid me
9. I enjoy avoiding arguments
10. Arguing over controversial issues improves my intelligence
11. I enjoy a good argument over a controversial issue
12. I try to avoid getting into arguments

Likelihood of giving opinion (used in Study 2)

(scale anchors: 1 (not at all likely) to 7 (very likely))

How likely would you be to raise your hand to offer your opinion?

Perceptions of confederate opinion (Manipulation check, Study 3)

(semantic differential, 7-point scale)

Now we'd like to know what you think your partner's political views are:

Very liberal:Very conservative

Perceptions of victimization (used in Study 1)

(responded on a slider, ranging from 0-100%)

Please estimate the percentage of each group that you think has been the victim of discrimination based on their group membership during their time at UCSB:

Women, men, Blacks, Latinos, Asians, Native Americans, Whites, Muslims, Jews, Christians, LGBTQ

Political ideology (used in Studies 1, 2, & 3)

(responded using semantic differential items)

My political positions are:

1 (strongly in favor of Democrats) to 7 (strongly in favor of Republicans)

1 (extremely liberal) to 7(extremely conservative)

Support for Censorship (used in Study 1)

Responded using check boxes (answer options: Yes, No, Not sure)

Lately there has been discussion at UCSB about which speakers are appropriate to bring to campus. Below is a list of topics that could be discussed at a lecture. Would you support UCSB allowing the following lectures to take place?

A presenter who argues that:

The US government should provide health insurance for everyone

The UN should create a global currency

Homosexuality is a disease that can be cured through conversion therapy

There should be a constitutional amendment banning the Electoral College and making president elected by popular vote

Transgenderism is a form of mental illness

Genetically modified foods are harmful to health

The FDA should more closely regulate nutrition labels

The more people who have guns, the lower the crime rate will be

Alt-right ideology

The earth's temperatures naturally ebb and flow, and climate change is not caused by humans

Willingness to Self-Censor (used in Studies 2 & 3)

(scale anchors: 1 (strongly disagree) to 5 (Strongly agree))

1. It is difficult for me to express my opinion if I others won't agree with what I say.
2. There have been many times when I have thought others around me were wrong but I didn't let them know.
3. When I disagree with others, I'd rather go along with them than argue about it.
4. It is easy for me to express my opinion around others who I think will disagree with me. (reverse coded)
5. I'd feel uncomfortable if someone asked my opinion and I know that he or she wouldn't agree with me.
6. I tend to speak my opinion only around friends or other people I trust.
7. It is safer to keep quiet than publicly speak an opinion that you know most others don't share.
8. If I disagree with others, I have no problem letter them know it. (reverse)

Appendix B: Meaning Extraction from Email Text

Category	Word	Loading	Example context
Discrimination	Homophobic	0.89	"...There is a difference between having different political views, and being outright **Racist**, **misogynistic**, **homophobic,** etc."
	Misogyny	0.89	
	Racism	0.79	
Deliberation	Understand	0.55	"Plus, how can you effectively argue against something you disagree with if you don't **understand** or refuse to listen to the other **side**'s argument?"
	Side	0.51	
	Voice	0.44	"There needs to be a **space** in which those who feel threatened can go and have their **voices** heard"
	Space	0.42	
	Hear	0.42	"…**hearing** opinions you don't agree with is healthy"
	Share	0.36	"Therefore no one should be banned from being able to publicly **share** their **opinion.** "
	Opinion	0.39	
	Disagree	0.39	"...they should be able to speak even if a majority of students **disagree** with their views"
	Oppose	0.35	"It provokes thought and healthy debate to have speakers that **oppose** liberal views…."
	Belief	0.34	"...individuals feel comfortable and free to express their beliefs"
	Negative	-0.31	"…by not allowing these **negative** speakers to come talk on campus…."
Presentation	Topic	0.56	"Going to college and learning new things does mean learning about **controversial topics**"
	Controversial	0.48	
	Issue	0.46	"….in order to engage in a well-rounded conversation regarding the **issue...**"
	Give	0.40	"…I still believe that speakers should be allowed to **give** presentations on college campuses"
	Education	0.37	"...preventing opposing viewpoints from being shared on campus would essentially be undermining the reason why higher **education** exists"
	Attend	0.31	"….they do not have to **attend** if they do not want to"
	Opportunity	0.32	"All individuals should have the ability to voice their opinions and have an **opportunity** to speak about difficult topics"
	Live	-0.34	"….well-rounded scholars that can learn to **live** in the outside world"

Safe space	Safe	0.47	"Milo is trying to have other people's rights taken away for something that they can not control, so why should we allow him to say those things that threaten others **safety** and well-being?"
	Marginalize	0.45	"If the speaker is clearly against a certain **marginalized** group and isn't even willing to address an opposing opinion, they are far too stubborn to be welcomed on campus"
	Learn	-0.38	"...as we can **learn** many new **perspectives** that we didn't hold before"
	Perspective	-0.41	
Diversity of Opinion	Protect	0.42	"….my understanding of the First Amendment is that speech, even hateful speech, is **protected** on public campuses"
	Liberal	0.42	"In times like these at a **liberal** university, it is becoming more and more awkward, uncomfortable, and downright scary to express anything that might be misconstrued as offensive or targeting a **minority**"
	Minority	0.31	
	College	-0.55	"**College campuses** are definitely supposed to be a safe place for people to express their opinions"
	Campus	-0.38	
	Create	-0.33	"Diversity of opinion benefits everyone and **creates** a safe environment for students…"
	Bring	-0.42	"The thing is, that the UC is an institution with their own values and if their **bring** speakers whit discriminating or racist ideals, that is what they are showing to their communities"
Campus life	Agree	0.50	"I **agree** that diversity on campus benefits everyone but along with diversity comes different opinions"
	School	0.32	"….when an university allows **speakers** to say offensive opinions in a campus, it may make some groups of people in the **school** to feel uncomfortable"
	Speaker	0.36	
	Comfortable	0.46	"College campuses should protect all students and provide a **comfortable** environment"
	Free	-0.65	"….you are preventing them from practicing the fundamental right of **free speech**"
	Speech	-0.48	
Role of University	Allow	0.49	"By not **allowing** these speakers to come to campus, these marginalized groups cannot be challenged in ways that they would be in more comfortable settings"
	Fact	0.37	"However, what is concerning is the **fact** that sometimes, the core values of certain groups are to demean marginalized groups or make them feel inferior"
	Hate	0.38	"If an individual feels that they are being personally attacked, they may be afraid of speaking up against the **hate** speech and be silenced as a result"
	Political	0.36	"By attacking people who express sentiments that are not **politically** correct, you are no better than anyone else"

Society	0.35	"I do believe that political diversity is very important to advocate because it would benefit **society** to be open-minded"
Attack	0.35	"Speakers who **attack** or demean certain groups of people should not be allowed to speak on college campuses"
University	0.52	"...**universities** should gauge whether or not the discussion will **benefit** the campus."
Benefit	0.31	

Note: Examples of participant texts are exactly how they were written, and were not edited for spelling or grammar.