

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

2021

THE BACKFIRING EFFECTS OF FORCED POLITICAL ALLEGIANCE IN THE ONLINE WORLD: THE CASE OF PREFERRED PRONOUNS ON TWITTER

Yin Lok Linus Chan

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Chan, Yin Lok Linus, "THE BACKFIRING EFFECTS OF FORCED POLITICAL ALLEGIANCE IN THE ONLINE WORLD: THE CASE OF PREFERRED PRONOUNS ON TWITTER" (2021). *Graduate Student Theses, Dissertations, & Professional Papers*. 11697.

<https://scholarworks.umt.edu/etd/11697>

This Dissertation is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

THE BACKFIRING EFFECTS OF FORCED POLITICAL ALLEGIANCE IN THE
ONLINE WORLD: THE CASE OF PREFERRED PRONOUNS ON TWITTER

By

YIN LOK LINUS CHAN

M.A. Psychology, University of Montana, Missoula, MT, 2019

B.A. Psychology, Trinity University, San Antonio, TX, 2017

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Psychology

The University of Montana
Missoula, MT

May 2021

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Lucian Gideon Conway, Chair
Psychology

Daniel Denis
Psychology

Stephen Yoshimura
Communications

Rachel Severson
Psychology

Allen Szalda-Petree
Psychology

Abstract

Political divisions in the United States have led to conflating political issues with political parties. People who possess a set of intercorrelated beliefs (e.g. pro-immigration, pro-choice) are assumed to be liberal, while those who hold opposing intercorrelated beliefs are assumed to be conservative. Impression management research suggests that in the online world, minor cues on one's social media profile that display political beliefs serve as indicators of people's political ideology. While people are free to declare their political views on personal social media accounts, what happens when people perceive that they are forced to appear aligned with one political group (whether they are a member of that group or not)? To examine this question, I had Twitter users respond to a scenario about preferred pronouns. Specifically, participants read a scenario where all Twitter users are either required to post cues on their Twitter profiles that are associated with political liberalism (or not). Results show that forced display of preferred pronouns increases public agreement overall but decreases private agreement for those who did not already post pronouns. Pressure to post preferred pronouns also indirectly decreased LGBT attitudes and increased concern for political correctness. These indirect effects were mediated by two well-known psychological phenomena that other work implicates in backfiring: Reactance and informational contamination. I conclude by discussing the implications of forced political allegiance in the online world.

Acknowledgements

As this dissertation represents my cumulative growth during my four years in graduate school, I express my gratitude and fondness to those who have supported me throughout this journey. These include professors, mentors, friends, and family, neatly separated into seven categories.

First, I thank my graduate advisor Dr. Luke Conway for training me to become a competent researcher and scholar. I always value your input when discussing research ideas, your mentorship with publishing papers, and your expertise in political psychology.

Second, I thank my lab mates - James McFarland, Shailee Woodard, Alivia Zubrod - for the adventures in Montana and the mutual support as we progressed the Ph.D. program together.

Third, my dissertation committee. Dr. Rachel Severson, thank you for your unwavering support and guidance on pedagogy. Dr. Stephen Yoshimura, thank you for serving on all three of my committees (master's thesis, comprehensive exam, doctoral dissertation) and providing valuable feedback from a communications perspective. Dr. Allen Szalda-Petree, thank you for leading the psychology department so well as department chair. Dr. Daniel Denis, thank you for teaching me everything about statistics while I was your teaching assistant for both undergraduate and graduate statistics.

Fourth, my friends. As inferential statistics tell us that we can use a sample to infer the population, I draw from my population of friends a sample of 26 important ones (I chose 26 to match my age when this is written). These friends (i.e. independent variables) have contributed to Linus' happiness and swagger (i.e. dependent variables). To avoid showing any more favoritism, I list these friends in alphabetical order by first name:

Aaron Fisher, Aaron Zhuo, Adrienne Lee, Alex Tobias, Anthony Chan, Braden Sides, Caleb Ho, Christopher Yu, David Herman, Harry Hughes, James Lee, Janice Wong, Jason Chen, Jason Lee, Jason Tam, Jeon Jungkook, Laris Li, Lorraine Lau, Maple Tang, Marvin Kwok, Matthew Schwark, Minnie Fu, Ronald Yim, Tessa Hughes, Tom Cheng, William Huang.

Fifth, professors from my undergrad institution, Trinity University. Dr. Kevin McIntyre, thank you for convincing me to major in psychology back in 2014 when I was a confused and undecided teenager. Dr. Harry Wallace, thank you for building my social psychology foundation; I used material that you taught in class and repackaged it in my class when I taught the same course. Dr. Carol Yoder, thank you for letting me conduct experimental psychology research in your lab. Somehow I liked research a bit too much which led me to pursue a doctoral program.

Sixth, my mentors. Dr. Cecilia Cheng, thank you for inspiring me to pursue cyberpsychology as my major research area; I always enjoy collaborating with you. Dr. Cory Clark, thank you for your advice on how to best communicate social psychology to the public. Dr. Jonathan Haidt, thank you for discussing your research on moral and political psychology with me over the years.

Finally, my family. I thank my dad Hubert, the original Dr. Chan, for providing opportunities to apply my research to non-academic arenas. I thank my mom Josephine for being the only family member who actually reads every single one of my publications from start to finish. And lastly, I thank my only sibling Ernest for supporting my numerous unconventional hobbies, which I will not name here.

The Backfiring Effects of Forced Political Allegiance in the Online World: The Case of Preferred Pronouns on Twitter

People have disagreed on countless political issues in every society. Virtually every large-scale issue that involves cooperation, conflict, or control is arguably political (Susskind, 2018). Hence, political issues such as education, healthcare, rights for minority groups, taxation, suffrage, justice, and democracy have been subject to alterations and reforms throughout history (Pinker, 2018). Political disagreements can sometimes be solved through discussion and debate, but sometimes when opposing ideologies collide, the dominant authority figure or group heavy-handedly forces their political views into society. Doing so often leads the populace to publicly display their political allegiance to that dominant view or risk punishment. As a result, there is a superficial display of public unity regardless of what people believed privately.

These instances of forced political allegiance occurred quite frequently in recent history before the Internet era. During the Chinese cultural revolution, every Chinese citizen had to own and carry Mao's little red book and display public agreement with Mao's ideas, one of which involved defying professors and intellectuals (Chan et al., 2009). During the Khmer Rouge regime, Pol Pot radically reformed society; there were affirmative action policies that recruited poor students into schools to learn radical socialist ideology. Those who displayed any deviance to the dominant ideology were killed (Clayton, 1998). In other communist regimes, there were bureaucratic regulations of what parents could name their child, such that Christian names were forbidden. Anti-communist ideas were also forbidden, to the point where hundreds of Czechoslovak historians were expelled from the universities and their works

disappeared from bookstores and libraries (Shore, 1996). During the Nazi regime, people who chose to hide Jews from Nazis were killed, which led to people in Germany publicly agreeing that they were not hiding Jews even when they privately helped Jews themselves (Blaustein, 2018). Regardless of what dominant ideology was present in what country, the populace had to publicly appear to agree with the dominant ideology or risk being punished.

Forced political allegiance occurs even in today's digital age where millions of people around the world are connected via the Internet and social media (Cheng, Chan, & Chan, 2021). Even in 2020, Hong Kong's pro-democracy protests have sparked a controversial national security law that permits local authorities to arrest anyone who shows dissent or subversion against Mainland China (Siu & Lau, 2020). While this legislation partly aims to restore order to Hong Kong by reducing protests that have caused the local economy to plummet, it is unclear what counts as dissent. It is possible that any speech or text on social media could be considered as thought crime if it is anti-Mainland China. Thus, Hong Kong citizens have to agree with the dominant ideology publicly both offline and online or risk punishment.

These cases of publicly displaying one's political allegiance despite one's private beliefs is a case of preference falsification (Kuran, 1987). Widespread public support may manifest on the surface, but the superficial consensus does not actually reflect the citizen's private thoughts. In other words, many falsely declare their political allegiance to avoid getting punished. Forced political allegiance goes against liberal ideals (e.g. freedom of expression, freedom of speech, freedom of press), and while the world overall has shown more adherence to liberal ideals for the past several decades worldwide (i.e.

there is more freedom, less sexism, racism, homophobia, more peace, prosperity, and equal rights; Pinker, 2018; Charlesworth & Banaji, 2019; Rosling et al., 2019), forced political allegiance may still exist in many forms – including in today’s online world.

Hence, this dissertation examined the underlying psychology of forced political allegiance in the online world. Specifically, I examined one social media platform that discusses politics on a daily basis: Twitter. The political issue I focused on is a recent political issue in the United States: declaring one’s own and addressing others by their preferred pronouns. Of course, I am not in any way equating the preferred pronoun issue with regimes and suppression of free speech mentioned above, but rather I am using this issue to study the effects of forced political allegiance. Independent of the positivity of any particular movement, the processes involved can backfire. Indeed, it is especially important to study those processes for positive movements – such movements would increase their likelihood of succeeding by understanding the processes involved in forced consensus. It is those processes I am investigating. Examining the forced political allegiance of the preferred pronoun issue on Twitter would thus serve as a useful case study to answer my primary research questions, which are: (1) Does forcing people to display their political allegiance in the online world backfire? (2) What are the underlying psychological mechanisms that explain the backfiring effect?

To evaluate these questions, I discuss political tensions in the United States, preferred pronouns as a recent manifestation of political impression management, and the potential backfiring effects of forced declaration of preferred pronouns.

Political Division in the United States Results in Conflating Political Issues with Political Parties

The United States today is highly polarized (Schaeffer, 2020). Politics in the United States have been divisive partly because of the mutually-exclusive two-party-system that tends to align political issues with political parties. Political ideologies tend to turn into biased groups where people assert their ingroup opinions as morally superior and the opposing group as morally inferior (Clark et al., 2019; Harris & Van Bavel, 2020). As a result, many political issues such as LGBT rights, immigration, abortion, healthcare, death penalty, universal basic income, and alcohol and drug use are politicized such that liberals tend to support one way and conservatives support another way. Even the COVID-19 pandemic has been framed as a liberal vs. conservative issue in terms of whether people should be wearing masks (Schneider, 2020; Malloy & Schwartz, 2020). These differences in political views can be explained by psychological theories such as moral foundations theory (Haidt, 2012), social dominance orientation (Duckitt & Sibley, 2016), threat avoidance (Crawford, 2017), governmental power over citizens (Chan & Conway, 2018), socio-ecological stressors (Conway, Chan, & Woodard, 2019), and ideological matching (Conway et al., under review).

Regardless of the causes of the political divide, there appears to be a general consensus that Republicans and Democrats disagree on almost all major political issues today (Fractured Nation, 2019). The disagreement lets people predict other's political ideology from their views on political issues. Data drawn from the Pew Research Center found that the majority of left-leaning individuals support illegal immigrants remaining in the U.S., support same-sex marriage, and support gun control. In contrast, the majority of

right-leaning individuals believe the opposite: they are against illegal immigrants remaining in the U.S., oppose same-sex marriage, and favor gun rights (Daniller, 2019; Heimlich, 2011a, 2011b).

In addition to disagreeing political viewpoints, there is a stark difference on how these two groups trust news outlets. Jurkowitz and colleagues (2020) found that Republicans were more likely than Democrats to rate most news sources as untrustworthy when news sources supported liberal views. In contrast, Democrats were much more likely to express that most news sources could be trusted (except for Fox news which supports conservative views). Not only is there disagreement about political issues, but also about who reports it as well.

But how do these disagreements manifest in today's online world? I argue that while today's digital age connects many Americans via social media, it enhances political divisions in the online world.

The Online World Augments Political Polarization

People's political views influence what news outlets they watch and how they react to the opposing political party. Strong partisans are more likely to selectively expose what media they see and are more likely to rate 'the media in general' as biased against their political party. While counterintuitive, this finding is explained by what news liberals and conservatives tend to watch (Barnidge et al., 2020). A liberal watching liberal news outlets would display little negative cues about those outlets, while simultaneously displaying negative cues about "the media." The same occurs for conservatives watching conservative news outlets.

Watching like-minded media often enhances ingroup favoritism and incites outgroup discrimination. Barnidge et al. (2020) recruited participants from Wisconsin - a politically divided state due to a bill that sparked mass protests - to complete measures on political opinion, media exposure, media partisanship, selective exposure, and perceived media bias. After controlling for basic demographics, results show that opinion extremity is positively associated with perceived media bias in general, and this relationship is mediated by selective exposure.

Polarization occurs even when people are exposed to media catered to the opposing political party. Levendusky (2013) found that news outlets often take a partisan stance to increase viewership, but doing so contributes to polarization. This is partly because encountering partisan news reminds people about their political allegiances, reinforces existing attitudes, and encourages confirmation bias. Levendusky (2013) had participants watch a series of real-world video clips that were left-wing (MSNBC), neutral (PBS news hour), or right-wing (Fox News), where all three clips were included and the order of videos shown was counterbalanced. Polarization was measured by how strongly participants agreed with their political party's clip relative to other clips. Results show that participants on both sides of the political spectrum rated their like-minded media with more extreme support. On a similar vein, Ashokkumar et al. (2020) found that laypeople are often willing to censor comments online that oppose their political beliefs regardless of whether the comment is offensive or not. These participants across both political parties actively censor content that challenges their political views, contributing to selective exposure and polarization.

But where did polarization come from in the online world? Social psychologist Jonathan Haidt explains three trends that contribute (Haidt & Rose-Stockwell, 2019). First, the introduction of the Facebook “like” and the Twitter “retweet” allowed content to be graded and ranked. Second, online posts that contain moral outrage (which stem from political disagreements) are more likely to be shared than mundane posts. Third, Facebook and Twitter present one’s newsfeed for optimal engagement, such that posts with the most likes, comments, and shares are presented instead of chronological order. Taken together, social media platforms facilitate the spread of outrage and this is exacerbated by mainstream media. To increase news engagement, journalists on social media platforms notice what is most popular and increase viewership of their mainstream media by sharing the same popular stories, which later gets shared back on social media. As a result, political stories that involve outrage receive more attention and have a higher likelihood of going viral.

Sharing outrage, some of which stems from political disagreements, could be both benevolent and malevolent. On one hand, more transparency on the Internet sheds light on unacceptable behavior, such as the prevalence of cover-ups of sexual assault in Hollywood (Farrow, 2019). On the other hand, sharing outrage contributes to a “call-out” culture where any post can be manipulated, taken out of context, or even fabricated to shame and ridicule. As such, political issues that benefit from debate are rarely communicated across parties publicly.

Political polarization in the online world contributes to preference falsification because people cannot always voice their true political views online. People who want to avoid being ridiculed for non-liberal views may publicly espouse attitudes in line with

political correctness (Boven, 2000). This is partly because publicly acknowledging a valid argument from someone in the opposing political party can easily get people in trouble from their own party, thereby leaving little room for compromise (Murray, 2019). For instance, Inbar and Lammers (2012) surveyed SPSP members (i.e. members who attend the largest annual social psychology conference in the world) and found that those who identified as conservatives rated academia as a more hostile climate to work in than liberals. Further, conservatives, unlike liberals, reported that they are afraid to divulge their political beliefs to colleagues. Though not a representative sample, many liberal respondents of Inbar and Lammers (2012)'s survey *self-reported* that they were inclined to discriminate against conservatives for paper reviews, grant reviews, symposium invites, and hiring decisions. These cases illustrate how some social psychologists publicly display some form of political allegiance regardless of what they privately believe, which may bias how social psychology research is conducted (Crawford & Jussim, 2018; Chan et al., 2018).

Preference falsification is further fueled by moral grandstanding (Grubbs et al., 2019), which occurs when a group who claims to represent the "correct" political view seeks to demonstrate moral superiority above other views. Moral grandstanding can take on many forms, such as public shaming of the opposite party, announcing that disagreements are signs of moral corruption, or exaggerating emotions to incite empathy when taking ideological positions (Bloom, 2017). Instances of moral grandstanding seek to publicly challenge or call-out political differences that can further lead to outrage that people often aim to avoid.

Because of political polarization and moral grandstanding, people are often motivated to carefully manage their online impression in a way that avoids public ridicule. This involves managing one's online impression in a public way that is widely accepted within a community, or concealing one's political views altogether.

Impression Management and Context Collapse in the Online World

In conventional face-to face-interactions, people are motivated to construct a favorable public image to maximize rewards and reduce punishments (Leary & Kowalski, 1990). Impression management includes two components: impression motivation and impression construction. Impression motivation is driven by one's goals to form a desired public image, while impression construction is the outcome of one's projected social image and self-concept (Leary & Kowalski, 1990).

While impression management research is often focused on traditional/face-to-face interactions, impression management theory is also applied in the online world. People are motivated to portray a favorable online image to those they encounter on their social networks, which includes careful displays of political views. Online impression management mainly concerns content-based management and network-based management (Walther et al. 2008). Content-based pertains to the contents a user places in the online world. As such, people carefully monitor what is placed online to appeal to those they want to impress. Network-based pertains to portraying one's involvement in certain groups, one's social role within each group, and how one communicates with others in that group. According to social psychologist Jonathan Haidt, people's sociometers - which are markers of self-esteem based on the degree of inclusion and exclusion in social groups - are broadcasted with public displays of friends, retweets, and

followers on various social media sites (Haidt & Rose-Stockwell, 2019). In other words, who one connects with on social networks influences how one's online profile is perceived (Walther et al., 2008).

The influence of social networks on the self has vast implications for impression management. This is because cyberspace allows multiple users and groups to interact simultaneously, and responses do not have to be synchronous. Different social media accounts encourage people to highlight different parts of themselves (e.g., LinkedIn vs. Facebook vs. Twitter). Hence, the audience across social media platforms is often conceptualized as 'imaginary' since users cannot be certain if they are currently being monitored and by how many people. Public social media accounts (e.g., public Twitter accounts) can be viewed by anyone anonymously with an Internet connection, which potentially allows people's online impressions to be under constant scrutiny. Since people cannot immediately react and cater to the audience, they may report feeling under constant surveillance. Ranzini and Hoek (2017) had Dutch participants report how strongly their behaviors are perceived to be influenced by those around them (measuring imaginary audience), and the degree to which they consciously portray a favorable self-image online (measuring online impression management). Results found that imaginary audience was a robust predictor for both content-based and network-based impression management.

Managing one's impression with the 'imaginary' audience is difficult because the nature of Internet connectivity makes a user connected to people from various social contexts (e.g. colleagues, family, friends, acquaintances) simultaneously and thereby collapses the context of different norms and different social groups into one broad

audience. This concept, known as 'context collapse,' forces people's self-presentation strategies to cater to various groups simultaneously. Groups that were previously detached in the offline world come together in the online world (Gil-Lopez et al., 2008). When one posts something sensitive online (e.g. a political opinion), they also have to consider that other recipients - including people who share their views and people who oppose their views - can see the post. Since any online post can be seen by potentially anyone in that particular social network (which may comprise of a range of political views), some people tend to be more apprehensive as more social spheres collide.

In some cases, people who are more apprehensive about their online impression management tend to increase their linguistic variability and self-disclose when it is universally appropriate. These people opt to portray a favorable image to their social networks and usually avoid discussing controversial or polarizing topics to avoid alienating their audience. Indeed, Marwick and Boyd (2010) surveyed Twitter users who have hundreds of thousands of followers and asked them who they tweet to, when they self-censor, who they imagine reading their tweets, and what makes them authentic. Results show that public tweets mix the personal and professional, and are suited for a general audience. Those Twitter users try to be authentic as best they can without too much self-disclosure.

In other cases, people are less apprehensive and publicly declare their opinions (e.g. political views) despite context collapse. Personal social media accounts allow people to freely and legally display their political views. For instance, some users put a rainbow flag to signal support for the LGBT community or liberal views more broadly. Other users put the "Make America Great Again" slogan on their profiles to signal

support for conservative views. Seemingly minor acts like these serve as cues for political impression management that allows people to infer other's political views from online profiles.

Political Impression Management in the Online World

Political impression management is important in the online world because communicating political views online reaches more people quickly than offline communication due to context collapse and the lack of latency when sharing information online. News stories received via tweets and retweets can come from any social group or acquaintance, such that recipients encounter political information even when not actively seeking them. Data from the General Social Survey conducted in 2004-2009 asked Americans to self-report their online news consumption and political ideology of their friends (Gentzkow & Shapiro, 2011). Results show that encountering someone online with opposing political views is about 45%, which is greater than encountering a friend with an opposing political view in person (35%). In other words, encountering an opposing political view is more likely online than offline.

Unfortunately, encountering someone else with the opposite political views online is often uncivil. In a nationally representative survey in 2016 conducted by the Pew Research Center, roughly two-thirds of Americans use social media, and about one-third engage in political activities online (Duggan & Smith, 2016). Interestingly, about 60% of participants described their interactions with opposing political views as frustrating, and 40% reported the tone on social media to be disrespectful for both liberals and conservatives. Incivility on Internet political discussions makes people perceive strong arguments as irrational and aversive (Popan et al., 2019). In their study, participants who

self-identified as liberal or conservative were asked to evaluate the rationality of the political outgroup's posts on a discussion board. These fabricated arguments were varied experimentally in two ways: civil or uncivil, and strong or weak argument. Results show that participants who saw uncivil content rated the comments as less rational regardless of argument strength (Popan et al., 2019). This suggests that there is plenty of animosity online that revolves around politics, which does not bode well for people who actively display their political allegiance online.

Given the hostility that emerges due to political differences, there are potential major consequences for revealing one's political views, be it discretely or blatantly. Because of the conflation of political issues with political parties, someone who voices support for political issues that aligns with a liberal view or views that are politically correct is assumed to be liberal who also supports liberal views such as globalism and pro-marijuana legalization, while someone who voices support that aligns with a conservative view (e.g. pro-life) is assumed to be conservative who also supports conservative views such as nationalism and anti-marijuana legalization. Minor cues on one's online profile such as a rainbow flag or a particular hashtag (e.g. #BlackLivesMatter, #MakeAmericaGreatAgain) are used to infer the person's political beliefs and the extent to which one abides by political correctness norms in the online world. Indeed, Landtsheer et al. (2008) described this phenomenon as "Perception politics" as it plays a major part in individuals' perception of politicians based on things such as their appearance and nonverbal behavior. As an example, consider an avid outdoorsman with pictures posing with firearms. According to Parker et al. (2019), 44% of Republicans are gun owners as compared 20% for Democrats. Knowing this, one may

view this outdoorsman's political stance as more conservative-leaning given the U.S.'s ongoing debate on gun rights and gun control politics. Conversely, if people make off-handed comments about their disdain for firearms, one may assume them to be more politically liberal. Another example would be the politicization of COVID-19 in the United States during the second half of 2020 whereby wearing a mask could be an indicator of liberalism and refusing to wear a mask could be an indication for conservative beliefs (Schneider, 2020).

While there has been research on how liberals and conservatives react to longstanding political issues such as LGBT rights, abortion, gun rights, death penalty (Barberá et al., 2015; Sterling et al., 2020), there is little research on a newer political issue that has a clear liberal vs. conservative divide: Preferred pronouns.

Preferred Pronouns as a Manifestation of Political Impression Management

In the past few years, a new political issue emerged into mainstream media: declaring and addressing people by their preferred pronouns. I want to stress again that the focus of this study is not to argue against this political position (or political correctness more broadly), but rather I use this issue as a cue for political allegiance and political correctness that may help us better understand the psychology of forced allegiance in the online world (see Conway et al., 2009; 2017, for discussions of how even good political aims can have unintended side effects).

Generally speaking, those who support the preferred pronoun issue (i.e. support declaring one's own pronouns and addressing others by preferred pronouns) tend to be liberal and supportive of political correctness, whereas those who oppose tend to be conservative (Dembroff & Wodak, 2018; Geiger & Graf, 2019). A Pew Research poll

found that the preferred pronoun issue is commonplace in the United States. Over one-third of Gen Z Americans (i.e. those born on 1995 or after) know someone who uses gender neutral pronouns, over one-quarter of millennials (i.e. those born between 1982 and 1994) know someone who uses gender neutral pronouns, and one-in-five Americans know someone who knows someone (i.e. second-degree of separation) who goes by a preferred pronoun. While the practice of preferred pronouns is gaining traction, the issue itself is divisive where people reported mixed support and opposition across all age groups (Geiger & Graf, 2019).

For much of the past few decades, the emphasis on people's preferred pronouns did not enter the mainstream political arena. However, according to Dennis Baron (2020), an English professor of the University of Illinois, the resurgence of interest in the debate on declaring one's preferred pronouns is due to a reinvigorating interest in gender politics surrounding transgendered rights. This newly formed discussion encourages more and more institutions and individuals to add their own preferred pronouns onto the ends of emails and social media accounts. Indeed, pronouns can have major legal implications as well as political ones. For example, Baron (2020) explains that for much of the past history the masculine pronoun "He" has been the generic and legal pronoun. However, this norm leaves out much of the population that does not fit a purely masculine form and thus there was a desire to challenge it and create a system that was more inclusive.

Broadcasting one's pronouns online is considered politically correct since it advocates for members of the sexual minority community by simultaneously signaling acceptance and reducing discrimination toward transgender and gender nonconforming individuals. Publicly declaring one's preferred pronouns gives trans and gender

nonconforming people a feeling of safety and belonging. A community that regularly puts pronouns online also normalizes discussions about gender (Wareham, 2019), which may be an effective way of reducing discrimination against sexual minorities.

However, in modern America, supporting preferred pronouns is likely not merely signaling support for gender politics, but also supporting liberal views in a larger sense. Because advocacy for LGBT persons is, in America, associated with being more liberal and politically correct (Swank et. al, 2013), it stands to reason that putting pronouns on Twitter profiles can make the Twitter community as a whole seem more liberal and politically correct.

Relevance to Today's Online World: Preferred Pronouns as a Recent Case of Political Allegiance on Twitter

Examining preferred pronouns serves as a recent case of political allegiance. An increasing number of people are putting their preferred pronouns on various online domains such as Emails, LinkedIn, and Twitter to signal support for LGBT individuals and to signal political allegiance to liberal principles (Pluckrose & Lindsay, 2020). This is evidenced by numerous Democratic candidates who ran for the 2020 Presidential election - Elizabeth Warren, Julian Castro, Pete Buttigieg, Cory Booker, Tom Steyer, and Bill de Blasio – who all have added preferred pronouns on their Twitter profiles to indicate inclusivity and solidarity among the LGBT community (Brammer, 2019; Soh, 2020). Indeed, Democratic Vice-President Kamala Harris included preferred pronouns on her Twitter profile.

I chose to examine Twitter as the social media platform to focus on because it is the platform where most political public discussions/disagreements occur on a daily

basis (relative to other platforms such as Facebook or LinkedIn). Many Twitter users already actively choose to declare their preferred pronouns as a sign of political allegiance - as evidenced by the Twitter bios of Democratic politicians, sexual minorities, LGBT allies, and transgender individuals (Saad, 2020; Shrier, 2020) - but what would happen if Twitter enacted a policy requiring it? No research to my knowledge examines the effects of forcing people to declare one's preferred pronouns and the potential backfiring effects, likely because of its recent emergence to the mainstream cultural narrative relative to other political issues.

People who choose to declare preferred pronouns on their own accord also tend to support addressing other people by their preferred pronouns. However, declaring one's pronouns and addressing others with their preferred pronouns carries a key distinction. People who do not declare one's own preferred pronoun to avoid appearing liberal publicly (or because they disagree with this issue) may still address others by their preferred pronoun as a sign of respect. Indeed, three of the most famous critics of the pronoun issue - Jordan Peterson, Debra Soh, Ben Shapiro - have publicly said that they are willing to address people by their preferred pronouns privately if asked to do so respectfully, but refuse to declare their own preferred pronouns (Paikin, 2016; The Daily Wire, 2019; Soh, 2020). Moreover, declaring one's preferred pronouns on one's profile signals political allegiance regardless of what one Tweets about, while those who address others' preferred pronoun may avoid appearing liberal by Tweeting about ideas and events instead of people. Hence, I focus on declaring one's own preferred pronoun instead of addressing others by their preferred pronouns.

Potential Backfiring Effects of Forced Declaration of Preferred Pronouns

Regardless of the political direction, forcing people to show political allegiance one way or another may have beneficial purposes, but forced consensus of political impression management has to be used cautiously as it can backfire. Research on the psychology of forced consensus has found that pressure to form a consensus (via top-down pressure) works to form a public consensus, but ultimately backfires beneath the surface privately (Conway & Schaller, 2005; Conway et al., 2009; Conway & Repke, 2019). For example, in experimental manipulations, people who thought about conscription (i.e. being forced to fight for their own country) reported a decrease of patriotism (i.e. love for one's own country) relative to people who were presented with the option of voluntarily joining the army (Chan & Conway, under review a¹). Similarly, people who imagined being forced to support an unfavorable applicant did so publicly but not privately (Chan & Conway, under review b). Further, people who thought about restrictive political correctness (PC) norms (i.e. public pressure that regulates one's speech) reported higher support for highly anti-PC Donald Trump relative to those who did not think about PC norms (Conway et al., 2017). Finally, people who thought about being pressured by a powerful authority figure to engage in pro-environmental behaviors reported less support for sustainable behaviors (Conway & Repke, 2019). In all these cases, the backfiring effects of forced consensus were attributed to two psychological mechanisms: reactance (emotional) and informational contamination (cognitive).

¹ Although this conscription's effect on patriotism paper is not yet published, it did win the International Council of Psychologist's early career research award in 2020.

Reactance

When people's freedom is threatened or removed (e.g. the choice to not fight in the army), people are motivated to restore their freedom by acting against the perceived pressure. Psychologists have long called this phenomenon reactance. Perceived pressure induces reactance on an individual level, and may consequently manifest in behaviors to resist the pressure and reestablish freedom. As an example, movie ratings that prohibit adolescents from watching certain films can backfire as those very same ratings attract adolescents due to reactance (Varava & Quick, 2015).

Prior research found that reactance is a psychological mediator that explains backfiring effects of forced consensus. Censorship induces reactance by removing the individual's freedom to not fight for one's country, and this reactance in turn led to decreased patriotism (Chan & Conway, under review a). People who thought about PC norms experience reactance as PC norms regulate what one can say, and this reactance in turn led to more support for anti-PC Donald Trump (Conway et al., 2017).

In the case of preferred pronouns, reactance will likely occur if people lose their choice of displaying (or concealing) their political allegiance. Under an impression management lens, forced display of one's preferred pronouns compromises impression construction as people can no longer fully control their desired online persona (Leary & Kowalski, 1990; Kramer & Haferkamp, 2011). As such, people's compromised impression construction may induce reactance. While some liberals may experience reactance despite agreeing with the pronoun issue, reactance may hit especially hard for those who do not agree with the pronoun issue or with liberal ideology.

While reactance as a psychological mechanism does not alone determine behavior because people may still comply and publicly agree with the policy at hand, people may still feel reactance privately and become motivated to act against the pressure when given the opportunity. Hence, while public pressure may manifest superficially on the surface, reactance could grow privately and sabotage the agreement that the pressure meant to create in the first place.

Informational Contamination

Informational contamination is the process where people discount shared information because they believe 1) the consensus is artificially created, 2) the information is inauthentic, or 3) the consensus is reflective of some political agenda (Conway & Schaller, 2005; Conway & Repke, 2019). Informational contamination will likely be elicited when the consensus seems to be constructed by top-down pressure (such as an authority figure's command) instead of reflecting the genuine beliefs of the persons comprising the consensus. Importantly, informational contamination frequently manifests in the same ways in the online world.

Prior research found that informational contamination is another psychological mediator that explains backfiring effects of forced consensus. Conscriptio induces informational contamination because conscription creates an artificial consensus where everyone is fighting for their country because they are legally required to, which in turn led to decreased patriotism (Chan & Conway, under review a). PC norms induce informational contamination because people attributed others' filter in language due to coercion instead of volition, which in turn led to increased support for anti-PC Donald Trump (Conway et al., 2017).

In the case of preferred pronouns, forcing every Twitter user to declare their preferred pronouns could backfire because people will attribute the declaration of preferred pronouns as an agenda for the Twitter community to appear liberal and support LGBT persons, such that people who abide by the new policy are doing it to stay on Twitter instead of one's volition to appear liberal. Since the consensus is contaminated, no one can be sure of other Twitter users' true political views as everyone on Twitter is forced to appear liberal and politically correct. Hence, I predict that everyone, regardless of whether they support the new policy, will experience informational contamination when forced to declare preferred pronouns. Informational contamination would in turn decrease support for the new policy, and possibly also influence one's support for LGBT persons and political correctness more broadly.

That being said, while informational contamination as a psychological mechanism does not alone determine behavior as people may very well comply and do what is forced publicly, people will likely not trust the consensus privately because they see consensus as a result of forced behavior instead of volitional behavior. Across scenarios of forced consensus, people may perceive behavior under an authority figure as attributions of obedience instead of the person's real beliefs.

Informational Contamination and Reactance in Preferred Pronouns

Both reactance and informational contamination could explain the effects of forced political allegiance. Forcing people to declare their preferred pronouns may cause an apparent public consensus in the short-term (since everyone declares their preferred pronouns and sees every other Twitter users with preferred pronouns), but this consensus may collapse in the long-term because it decreases private agreement as well as

decreased support for LGBT persons and increased concern for political correctness more broadly. To the degree that this is true, any pressure to make social media users to lean politically in one way or another should be used cautiously.

Reducing Backfiring Effects of Forced Consensus

So what can be done? Is it possible to change social norms without inducing backfiring? Forcing people with laws and rules is sometimes necessary for society to progress, but often times forced consensus comes with a cost. Given that forced consensus tends to backfire in the long-term due to reactance and informational contamination, it is useful to examine how to instill a forced consensus that works in the long-term. One solution proposed by Conway et al. (2021) is to convince the population that one view is normative. For instance, slavery was a divisive issue in America but no longer is because one side became widely (and rightfully) accepted.

In the context of preferred pronouns, creating long-lasting consensus means persuading Twitter users that posting pronouns is *genuinely* (and not artificially) normative. One way to achieve this is to inform Twitter users that there is a consensus on Twitter to post one's preferred pronouns regardless of political ideology. Since past research suggests that people do not make accurate estimations regarding the true consensus on who supports political correctness policies (Boven, 2000; Fingerhut, 2016), it may help alleviate informational contamination when participants are presented with information that suggests a majority consensus about posting preferred pronouns on Twitter.

The Current Research

In the present study, I examined how participants react to a hypothetical new policy on Twitter about preferred pronouns. The policy had two components: 1) participants will be forced to declare one's preferred pronouns (or not), and 2) participants will be told that there is a consensus that Twitter users agree with posting preferred pronouns for both liberals and conservatives (or not). I manipulated whether participants were pressured to support the policy and whether participants were told about the consensus, and asked participants to report their public and private agreement towards the policy, along with attitudes about LGBT persons and concern for political correctness. I then measure if this pressure backfires via reactance and informational contamination. I also asked participants to report if they already post preferred pronouns on their own Twitter bio, and used that measure as a grouping variable to see if that alters mean patterns.

Hypotheses

Since I predicted different mean patterns for the DVs [public agreement, private agreement, support for LGBT persons, concern for political correctness], I discuss these hypotheses separately.

Hypotheses Related to Public Agreement

H1: There will be a main effect of forced declaration on one's own preferred pronoun, such that it increases public agreement with the policy.

H2: There will be a declaration of one's own preferred pronouns x consensus interaction, such that declaring one's pronouns increases agreement if there is consensus, but shows a weaker effect without consensus.

Hypotheses Related to Private Agreement

H3: There will be a main effect of forced declaration one's own preferred pronoun, such that it decreases private agreement with the policy.

H3a: Forced declaration one's own preferred pronoun will decrease private agreement (H3) indirectly via reactance and informational contamination.

H4: There will be a forced declaration of one's own preferred pronouns x consensus interaction, such that declaring one's pronouns decreases private agreement if there is consensus, but shows a weaker effect without consensus.

Hypotheses Related to Attitudes Towards LGBT Persons

H5: There will be a main effect of forced declaration on one's own preferred pronoun, such that it decreases attitudes towards LGBT persons.

H5a: Forced declaration of one's own preferred pronoun will decrease attitudes towards LGBT persons (H5) indirectly via reactance and informational contamination.

Hypotheses Related to Concern for Political Correctness.

H6: There will be a main effect of forced declaration one's own preferred pronoun, such that it increases concern for political correctness.

H6a: Forced declaration of one's own preferred pronoun will increase concern for political correctness (H6) indirectly via reactance and informational contamination.

Method

Participants

400 U.S. participants were recruited via Amazon's Mechanical Turk (MTurk). MTurk has previously been validated for research relevant to U.S. politics (see e.g., Clifford, Jewell, & Waggoner, 2015; Conway, Houck, Gornick, & Repke, 2017; Conway

& McFarland, 2019; Conway & Woodard, 2020) and it generally shows similar results as other samples (see, for example, Conway et al., 2017; Houck, Conway, & Repke, 2014).² Of the 400 participants, 10 were excluded because they failed the reading check question³, and an additional 130 were excluded because they failed the manipulation check item (described below). Thus the final sample was 260.

Within the final sample of 260 participants (56.5% male, 42% female, 1.5% unreported), 76.9% reported as Caucasian, and 77.3% of participants reported as heterosexual. Participants ranged from 18 to 75 years old ($M = 37.55$, $SD = 11.23$). The sample was slightly right-leaning politically (4.85 on a political conservatism scale with 4.5 as the midpoint). 112 out of the 260 participants reported that they currently post pronouns on their Twitter profile.

Inclusion Criteria

Participants were asked: "Do you have a personal Twitter account that you actively use?" scored dichotomously (yes vs. no). Participants who answered no were not permitted to complete the study.

Twitter Usage

Participants completed a series of questions pertaining to their Twitter usage adapted from Kang and Wei (2020), which was originally about Facebook usage. Questions asked how long ago the participant started using their personal Twitter account, number of accounts they follow, number of followers they have, the privacy setting of their account (dummy coded as private vs. public), frequency of tweeting,

² Given recently-identified potential issues with MTurk (e.g., Kennedy et al., 2019), I aimed to ensure the highest quality of data by including several screener questions that participants had to answer correctly to be included in the study. Evidence suggests that MTurk still produces excellent data given such safeguards (Kennedy et al., 2019).

³ The reading check question was a simple "This is a reading check. Please select 'agree.'"

frequency of retweeting, and how often they use Twitter to browse other profiles. The frequency of posting content online and browsing other profiles were scored on a 1-5 scale, while participants entered a specific number on followers and accounts following.

Independent Variables

Each participant read a scenario modeled after previous work on pressured agreement in other domains (Conway & Schaller, 2005; Conway et al., 2009; Conway & Repke, 2019; Chan & Conway, under review a; Chan & Conway, under review b). In these scenarios, they were asked to imagine that Twitter is enacting a new policy about preferred pronouns. The scenarios varied on two variables, and an additional grouping variable on whether participants already post preferred pronouns was included for analyses.

Forced Declaration of Own Preferred Pronouns Manipulation. In all scenarios, Twitter enacts a new policy that will come into effect in a couple of months. In the Pressure condition, participants were told that Twitter's new policy that supports diversity and inclusion requires that all users declare their preferred pronoun on their Twitter profile or their Twitter account will be suspended. In the No Pressure condition, participants were told that Twitter's new policy promotes user engagement and all users can customize their background theme, but this feature is entirely optional.

Consensus on Preferred Pronouns Manipulation. In the consensus condition, participants were told that Twitter conducted an anonymous poll that revealed a strong consensus where the majority of the Twitter community, which includes both liberals and conservatives, support posting preferred pronouns. In the no consensus condition,

participants were told that Twitter has no data about how the Twitter community feels about posting preferred pronouns.

All four versions of the scenarios in this 2 (Pressure) X 2 (Consensus) design are listed in Appendix A. All vignettes are approximately equal in length and each participant was randomly assigned to read one version.⁴ Participants were required to read their assigned vignette for at least 30 seconds before proceeding.

Dependent Variables

After participants read their assigned vignette, they were asked to complete the following measures in this order.

Public Agreement. Public agreement was measured with the following items adapted from prior work (Conway et al., 2009; 2017): “In the scenario, I would publicly comply with Twitter's new policy when it becomes effective in two months.” and “In this scenario, I would publicly endorse Twitter’s new policy” and “In this scenario, I would publicly post my preferred pronouns on my Twitter profile when the policy becomes effective” (1-7 scale, strongly disagree to strongly agree), with higher score indicating more public agreement. This scale demonstrated high reliability, $\alpha = .82$.

Private Agreement. Private agreement was measured with the following items adapted from prior work (Conway et al., 2009; 2017): “In this scenario, if my opinions about Twitter’s new policy were kept only to myself, I would willingly support Twitter’s new policy” and “In this scenario, I would privately support Twitter’s new policy” and “In this scenario, I would privately support posting my preferred pronouns on my Twitter

⁴ I used vignettes in this study because there are plenty of variables that influence people’s agreement with putting preferred pronouns on Twitter, and using vignettes isolates these variables to just the manipulations of pressure and consensus. Despite the drawbacks associated with using hypothetical scenarios, using vignettes and self-reports is still a useful way to study people’s political opinions in social psychology research.

profile.” (1-7 scale, strongly disagree to strongly agree), with higher score indicating more private agreement. This scale demonstrated high reliability, $\alpha = .91$.

Attitude Towards LGBT Persons Scale. Attitude towards LGBT persons was measured by the 9-item attitude subscale of the LGBT assessment survey (Logie et al., 2007). Sample items include “If a person feels that they belong to a different gender than the one they were born into, they should do everything to overcome these feelings“ and “Bisexuality is merely a different kind of lifestyle that should not be condemned.” Participants responded on a 5-point scale (1 = strongly disagree, 5 = strongly agree), with higher scores indicating more positive attitudes towards LGBT persons. The original published subscale demonstrated high reliability, $\alpha > .90$, and the scale demonstrated high reliability in my sample, $\alpha = .92$. The full scale is in Appendix B.

Concern for Political Correctness. Concern for political correctness was measured by one item from the American National Election Studies (ANES, 2020) adapted for this scenario. The item was “Some complain that too many people are easily offended these days and are too quick to police what others say. They refer to this as “political correctness”. Do you think that political correctness would be a serious problem when considering Twitter’s new policy” Participants responded on a 5-point scale (1 = not serious at all, 5 = extremely serious), with higher scores indicating greater concern for political correctness.

Proposed Mediators

Informational Contamination. Participants completed a 3-item scale that measured their informational contamination in the context of Twitter’s hypothetical new policy (adapted from Conway & Schaller, 2005; Conway et al., 2009; Conway &

Repke, 2019; Chan & Conway, under review a; Chan & Conway, under review b). Items were “In the scenario, I believe that Twitter’s new policy is part of a scheme to make Twitter users appear more liberal” and “In the scenario, I would distrust Twitter’s new policy because I assume it is reflective of some political agenda” and “In the scenario, I would distrust new policies like these because I assume it is reflective of some agenda.” Participants responded on a 7-point scale (1 = strongly disagree, 7 = strongly agree), with higher scores indicating greater informational contamination. This scale demonstrated high reliability, $\alpha = .88$.

Reactance. Participants completed a 3-item scale that measured their reactance in the context of Twitter’s hypothetical new policy (adapted from Conway & Schaller, 2005; Conway et al., 2009; Conway & Repke, 2019; Chan & Conway, under review a; Chan & Conway, under review b). Items were “In the scenario, I was aggravated by pressure I felt with regards to Twitter’s new policy” and “In the scenario, I felt that there was pressure to support Twitter’s new policy, which makes me want to not support the policy” and “In the scenario, other Twitter users’ support for Twitter’s new policy makes me feel as if I should not support the policy – just to show that I have the right to make up my own mind.” Participants responded on a 5-point scale (1 = strongly disagree, 5 = strongly agree), with higher scores indicating greater reactance. This scale demonstrated high reliability, $\alpha = .88$.

Manipulation Check

After participants read their assigned vignette and answered the scenario-specific dependent measures, participants answered the manipulation check item. This item was “In the scenario, did Twitter's new policy require that Twitter users post their preferred

pronouns on their Twitter profiles in a couple of months?” (yes vs. no). Participants who answered incorrectly (n = 130) were excluded from analyses.

Exploratory Moderators

Currently Posting Preferred Pronouns. Participants were asked if they currently post preferred pronouns on their own Twitter profile (yes vs. no).

Familiarity of Preferred Pronouns. Participants were asked how familiar they are with posting pronouns on social media. The item reads “How familiar are you with the concept of people posting their preferred pronouns (e.g. he/him, she/her, they/them) on social media?” and participants answered on a 1-9 scale (1 = unfamiliar, 9 = familiar). Participants were also given a box to qualitatively describe what posting pronouns online mean.

Pronouns as Appearing Liberal. Participants were asked if they think posting preferred pronouns will make them appear politically liberal. The item reads “I think posting preferred pronouns online would make others think I am: liberal/conservative” and participants answered on a 1-9 scale (1 = liberal, 9 = conservative).

Twitter Importance. Participants were asked how important Twitter means to them. The item reads “Twitter is very important to me” and participants answered on a 1-9 scale (1 = strongly disagree, 9 = strongly agree).

Political Ideology. Participants completed two standard items anchored by Liberal/Democrat at the low end and Conservative/Republican at the high end (e.g., Jost et al., 2008). These were averaged into a *Political Ideology* score, ranging from 1 to 9. The two items demonstrated high reliability, $\alpha = .84$.

Demographics

Finally, participants were asked to report their age, biological sex, gender identity (man, woman, non-binary), ethnicity, sexual orientation [Heterosexual/ Homosexual/ Bisexual/ Pansexual/Asexual/Other [with box to enter text]], education, and marital status.

Results

Given that participants who already post their own preferred pronouns on Twitter might perceive the hypothetical scenario differently than those who do not already post - since the scenario asks participants to imagine being forced to do something they already do - it is useful to conduct analyses on the data with the full sample of $N = 260$ with currently posting pronouns on Twitter [yes vs. no] as an additional IV, as well as analyses on the non-posting only sample of $N = 148$ participants who currently do not post preferred pronouns on Twitter. Indeed, both samples reveal meaningful results. Henceforth I refer to these two samples as the “full sample” and “non-posting only sample.”

Results for Full Sample

Primary Analyses

Separate 2 (Pressure to post pronoun: yes vs. no) x 2 (Consensus to post pronouns: yes vs. no) x 2 (Currently post pronouns: yes vs. no) factorial ANOVAs were conducted to examine the effect of the IVs on the dependent measures: public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness. Tables 1, 2, and 3 display the means for the four outcome variables for the full sample, posting-only sample, and non-posting only sample respectively.

In terms of public agreement, there was a main effect of currently posting pronouns, $F(1, 252) = 33.02, p < .001, \eta p^2 = .12$. Those who currently post pronouns reported more public agreement (currently post $M = 5.64$ vs. no post $M = 4.49, n = 260$). Inconsistent with H1 and H2, there were no effects of pressure or consensus, or any interaction (p 's $> .20$). Figure 1 displays means for public agreement for the full sample.

In terms of private agreement, there was main effect of currently posting pronouns, $F(1, 252) = 29.43, p < .001, \eta p^2 = .11$. Those who currently post pronouns reported more private agreement (currently post $M = 5.47$ vs. no post $M = 4.26, n = 260$). There was also a pressure x currently post interaction, $F(1, 252) = 6.88, p = .009, \eta p^2 = .03$, such that pressure increases private agreement for those who currently post pronouns, but pressure decreases private agreement for those who do not post. Inconsistent with H3 and H4, there was no main effect of pressure and no main effect of consensus, and no other interactions approached significance (p 's $> .22$). Figure 2 displays means for private agreement for the full sample.

To examine if pressure increased public agreement but decreased private agreement, a mixed ANOVA with agreement type (public vs. private) as the within-subjects variable and pressure (yes vs. no) as the between-subjects variable was conducted. Results show an interaction, $F(1, 258) = 8.56, p = .004, \eta p^2 = .04$, such that pressure increased public agreement but pressure decreased private agreement. Thus, although neither H1 or H3 was directly supported, this does suggest that pressure pulled public and private agreement in different directions – consistent with my original hypothesis.

In terms of attitudes towards LGBT persons, there was a main effect of pressure, $F(1, 252) = 5.59, p = .019, \eta p^2 = .02$. Consistent with H5, pressure decreased attitudes towards LGBT persons (Pressure $M = 4.52$ vs. No Pressure $M = 5.00, n = 260$). There was also a main effect of currently post pronoun, $F(1, 252) = 27.76, p < .001, \eta p^2 = .10$, such that those who currently post reported less favorable attitudes (Currently post $M = 4.23$ vs. No post $M = 5.29, n = 260$). There was also a pressure x currently post interaction, $F(1, 252) = 6.34, p = .012, \eta p^2 = .03$, such that pressure decreased LGBT attitudes only for those who currently post pronouns; pressure did not decrease LGBT attitudes for those who do not currently post. No other main effect of interactions reached significance (p 's $> .22$).

In terms of concerns for political correctness, there was a main effect of pressure, $F(1, 252) = 9.06, p = .003, \eta p^2 = .04$. Consistent with H6, pressure increased PC concerns (Pressure $M = 3.49$ vs. No Pressure $M = 2.30, n = 260$). There was also a main effect of currently post pronoun, $F(1, 252) = 10.55, p = .001, \eta p^2 = .04$, such that those who currently post reported increased PC concerns (Currently post $M = 3.51$ vs. No post $M = 2.95, n = 260$). No other main effect or interactions reached significance (p 's $> .07$).

Mean Patterns for Mediators

While reactance and informational contamination were quite strongly correlated ($r = .78, p < .001$), past research found that reactance and informational contamination would differentially affect the impact of pressure on outcome variables. Hence, separate 2 (Pressure to post pronouns: yes vs. no) x 2 (Consensus to post pronouns: yes vs. no) x 2 (Currently post pronouns on Twitter: yes vs. no) factorial ANOVAs were conducted with reactance and informational contamination as the DVs.

As expected, there was a main effect of pressure on reactance, $F(1, 252) = 23.87$, $p < .001$, $\eta^2 = .09$, such that increased pressure led to increased reactance (Pressure $M = 3.27$ vs. No Pressure $M = 2.47$, $n = 260$). There was also a main effect of posted pronouns on reactance, $F(1, 252) = 27.68$, $p < .001$, $\eta^2 = .10$, such that those who posted pronouns reported increased reactance (Currently post $M = 3.30$ vs. No post $M = 2.44$, $n = 260$). No other main effect or interactions reached significance (p 's $> .12$). Figure 3 displays means for reactance for the full sample.

Also as predicted, there was a main effect of pressure on informational contamination, $F(1, 252) = 14.51$, $p < .001$, $\eta^2 = .05$, such that pressure led to increased informational contamination (Pressure $M = 4.64$ vs. No Pressure $M = 3.73$, $n = 260$). There was also a main effect of posted pronouns, $F(1, 252) = 15.46$, $p < .001$, $\eta^2 = .06$, such that those who posted pronouns reported more informational contamination (Currently post $M = 4.65$ vs. No post $M = 3.72$, $n = 260$). Interestingly, there is a consensus x posted pronoun interaction, $F(1, 252) = 4.42$, $p = .036$, $\eta^2 = .02$, where consensus reduced informational contamination for those who already post pronouns, but consensus had no effect for those who do not post. No other main effect or interactions reached significance (p 's $> .09$). Figure 4 displays means for informational contamination for the full sample.

Mediation Analyses

To examine the indirect effects of the hypothesized $X \rightarrow M \rightarrow Y$ paths where the mediating variables were reactance and informational contamination, I followed recommended current practices for testing indirect effects. Specifically, I used the PROCESS macro (Hayes, 2018; model 4) to compute both normal tests of indirect effects

and bootstrapped confidence intervals (using 5000 samples) for each $X \rightarrow Y$ indirect effect with reactance or informational contamination as the mediator variable. In total, twenty-four separate mediation analyses were conducted to examine the effect of the three IVs (X) on the four DVs (Y) via two mediator variables (M). Importantly, the lack of direct $X \rightarrow Y$ relationships do not invalidate indirect effects (Darlington & Hayes, 2017). Mediation results for the full sample are depicted in Tables 4 and 5.

First, I examine the indirect effects of pressure on outcome variables. Consistent with H3a, pressure significantly increased both reactance and informational contamination, which in turn significantly decreased public agreement (indirect effects $p < .05$), but not private agreement (indirect effects $p > .05$). Consistent with H5a, there were indirect effects of pressure on LGBT attitudes, such that pressure's effect on both reactance and informational contamination in turn decreased LGBT attitudes (indirect effects $p's < .05$). Consistent with H6a, there were indirect effects of pressure on PC concerns, such that pressure's effect on both reactance and informational contamination in turn increased PC concerns (indirect effects $p's < .05$).

Next, I examined the indirect effects of consensus on outcome variables. There were no significant direct or indirect effects (all $p's > .05$).⁵

⁵I also examined the indirect effects of current pronoun posting on outcome variables. Results revealed that both reactance and informational contamination mediated the effects of posting pronouns on three outcome variables (i.e. all DVs but not private agreement). Specifically, both mediators reversed the effect of posting pronouns on public agreement. There was a direct effect that posting pronoun predicted increased public agreement, but indirect effects reveal that posting pronouns predicted increased reactance and informational contamination, which in turn decreased public agreement. Additionally, while there were no direct effects of posting pronouns on LGBT attitudes or PC concerns, indirect effects reveal that pressure indirectly decreased LGBT attitudes and increased PC concerns.

Moderation Analyses

To analyze whether our moderators (age, political orientation) influenced the main effect of pressure on the outcome variables, I followed standard current practices for testing the moderating effect of a continuous variable on the relationship between two other continuous variables via simultaneous regression (Hayes, 2018). Specifically, I used the PROCESS macro (Hayes, 2018; model 1) to examine if pressure's effects on outcome variables were altered at different levels of the moderator variables. Results revealed no moderating effect of age, but there was a pressure x political orientation interaction on public agreement, $p = .027$: In no pressure conditions, liberals agreed with the policy more than conservatives; however, the presence of pressure eliminated the effect of liberalism on agreement, such that political orientation no longer influenced agreement when pressure was introduced.

Results for Non-Posting Only Sample

Because the scenarios are catered to participants who have not already posted pronouns on Twitter, I also ran analyses only on those participants who had *not* already posted their preferred pronouns.⁶ This allowed me to answer the question: What happens when participants who do not currently post pronouns are forced to do so?

Primary Analyses

Separate 2 (Pressure to post pronoun: yes vs. no) x 2 (Consensus to post pronouns: yes vs. no) factorial ANOVAs were conducted to examine the effect of the IVs on the dependent measures: public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness.

⁶ While this second set of analyses were conducted on a subset of the full sample, I did not run Bonferroni corrections (or any other p-value adjustment) because the goal was to see if the same set of results replicated within this subset.

In terms of public agreement, there were no main effects or interactions (p 's > .72). Contrary to H1, pressure had no effect on public agreement ($p = .938$). Contrary to H2 and H4, consensus had no effect on public agreement ($p = .100$). Table 3 displays the means for the four outcome variables for the non-posting only sample. Figure 5 displays means for public agreement for the non pronoun posting sample.

However, in terms of private agreement, H3 was confirmed as there was a main effect of pressure, $F(1, 144) = 7.42, p = .007, \eta^2 = .05$, such that pressure decreased private agreement (Pressure $M = 3.84$ vs. no Pressure $M = 4.69, n = 148$). No other main effects or interactions reached significance (p 's > .34). Figure 6 displays means for private agreement for the non pronoun posting sample.

Importantly, as with the whole sample, a mixed-model ANOVA with agreement type (public vs. private) as the within-subjects variable and pressure (yes vs. no) as the between-subjects variable was conducted to examine if there was indeed a public vs. private difference. Results show an interaction, $F(1, 146) = 12.72, p < .001, \eta^2 = .08$, such that pressure has no effect on public agreement but has a strong negative effect on private agreement. This again suggests, consistent with my theoretical framework, that pressure pulls public and private agreement in different directions.

In terms of attitudes towards LGBT persons, there were no main effects or interactions (p 's > .77). Contrary to H5, pressure had no effect on LGBT attitudes ($p = .905$).

In terms of concern for political correctness, there were no main effects or interactions (p 's > .20). Contrary to H6, pressure had no effect on PC concerns ($p = .203$).

Mean Patterns for Mediators

Separate 2 (Pressure to post pronouns: yes vs. no) x 2 (Consensus to post pronouns: yes vs. no) factorial ANOVAs were conducted with reactance and informational contamination as the DVs.

When reactance is the DV, there was a main effect of pressure, $F(1, 144) = 16.28$, $p < .001$, $\eta p^2 = .10$, such that pressure increased reactance (Pressure $M = 2.85$ vs. no Pressure $M = 2.04$, $n = 148$). No other main effect or interaction reached significance (p 's $> .67$). Figure 7 displays means for reactance for the non pronoun posting sample.

When informational contamination is the DV, there was a main effect of pressure, $F(1, 144) = 6.27$, $p = .013$, $\eta p^2 = .05$, such that pressure increased informational contamination (Pressure $M = 4.09$ vs. no Pressure $M = 3.35$, $n = 148$). No other main effect or interaction reached significance (p 's $> .63$). Figure 8 displays means for informational contamination for the non pronoun posting sample.

Mediation Analyses

Like before, I used the PROCESS macro (Hayes, 2018; model 4) to compute both normal tests of indirect effects and bootstrapped confidence intervals (using 5000 samples) for each $X \rightarrow Y$ indirect effect with reactance or informational contamination as the mediator variable. In total, sixteen separate mediation analyses were conducted to examine the effect of the two IVs (X) on the four DVs (Y) via two mediator variables (M). Mediation results for the non-posting only sample are depicted in Tables 6 and 7.

Mediation analyses on this group of participants replicated the indirect effects conducted on the full sample. But this time, indirect effects of pressure were significant on all four on outcome variables (instead of three). Consistent with H3a, pressure

significantly increased both reactance and informational contamination, which in turn significantly decreased both public agreement (indirect effects $p < .05$) and private agreement (indirect effects $p < .05$). Consistent with H5a, pressure's effect on both reactance and informational contamination in turn decreased LGBT attitudes (indirect effects $p's < .05$). Consistent with H6a, pressure's effect on both reactance and informational contamination in turn increased PC concerns (indirect effects $p's < .05$).

There were no significant direct or indirect effects of consensus on outcome variables (all $p's > .05$).

Moderation Analyses

Like before, I followed standard current practices - PROCESS macro (Hayes, 2018; model 1) to test the moderating effect of age and political orientation on the relationship between pressure and outcome variables via simultaneous regression (Hayes, 2018). Results revealed no moderating effects, suggesting that the pressure manipulation's effect on the outcome variables was not differentially impacted by age or political orientation.

Examining the Characteristics of those who Currently Posted Pronouns on Twitter

Primary analyses reveal many robust main effects and interactions on how those who currently post pronouns (vs. those who do not post) view the hypothetical policy, LGBT attitudes, and PC concerns differently. But why is that? How do participants who currently post pronouns differ?

There are numerous significant differences between those who currently post pronouns compared to those who do not. Independent samples t-tests reveal that – using currently post vs. no post as the between-subjects variable – those who currently post

rated Twitter as more important to them ($t(258) = 11.17, p < .001$; post pronouns $M = 7.33$ vs. no post $M = 4.49, d = 1.5, n = 260$), are more familiar with the concept of posting pronouns on Twitter ($t(258) = 3.88, p < .001$; post pronouns $M = 7.37$ vs. no post $M = 6.63, d = 0.13, n = 260$), but also rated posting pronouns as appearing more conservative ($t(258) = 11.57, p < .001$, post pronouns $M = 6.23$ vs. no post $M = 3.07, d = 1.4, n = 260$), and rated themselves as more conservative ($t(258) = 7.06, p < .001$, post pronouns $M = 5.98$ vs. no post $M = 3.99, d = 0.89, n = 260$). Chi-square test revealed that only 34% of heterosexuals already post pronouns, while 77% of sexual minorities already post pronouns, $\chi^2(1) = 33.07, p < .001$.

Further Independent t-test and chi-square found no significant age differences, $t(258) = .60, p = .55$, or gender differences comparing males and females, $\chi^2(1) = 0.21, p = .65$, for those who post pronouns vs. those who do not post.

It is curious as to why those who currently post pronouns self-report as being more conservative and rate posting pronouns as appearing relatively more conservative. To better understand what is going on, I ran analyses that tried to disentangle the effects of conservatism from those of sexual orientation.

Disentangling Political Orientation and Sexual Orientation

Using the full sample, conservatism predicted increased private agreement when pressured to appear politically liberal. This is counterintuitive, but perhaps it can be explained by sexual orientation.

Compared to liberals, conservatives in this sample reported less favorable attitudes towards LGBT persons ($r = -.56, p < .001$) and conservatives reported increased PC concerns ($r = .55, p < .001$). These two correlations are not surprising, but what is

surprising is that those who identify as LGBT (i.e. non-heterosexuals) tend to rate themselves as more conservative. An independent samples t-test – using heterosexual vs. non-heterosexual as the IV – found that sexual minorities are more right-leaning in our sample, $t(258) = 4.11, p < .001, d = .62$ (heterosexual $M = 4.53$ vs. sexual minority $M = 6.00, n = 260$).

Conservatives in my sample reported less favorable views towards LGBT persons, but many conservatives in this same sample self-identified as LGBT. Since there is good reason to assume LGBT persons support LGBT rights, is posting pronouns on Twitter more influenced by conservatism or sexual orientation? To disentangle the two empirically, I examined whether the effect of political orientation on posting pronouns is moderated by sexual orientation, using PROCESS macro (Hayes, 2018; model 1). Results revealed a political orientation x sexual orientation interaction, $p = .016$, such that an increase in conservatism is associated with an increased likelihood of posting pronouns but only for sexual minorities. There is no relationship on conservatism and posting pronouns for heterosexuals.

These findings point to counterintuitive effects of those who post pronouns. It is unclear as to why pronoun-posters self-reported to be more conservative, and the variables within my data do not provide a clear explanation.

Discussion

Forced display of politically-loaded impression management cues has consequences. In line with expectations, the results provided support for hypothesized effects of the pressure with similar effect sizes from published research that used scenarios of forced consensus. However, contrary to expectations, the results provided no

support for the consensus manipulation. With the full sample, forcing people to post preferred pronouns resulted in increased public agreement, but no change in private agreement. With the non-posting only sample, forcing people to post preferred pronouns did not affect public agreement, but decreased private agreement. What explains this backfiring effect? Psychological mediators of forced consensus tell a rich story.

Effects of Reactance and Informational Contamination

Forced consensus elicits backfiring effects. Pressure predicted increased reactance and informational contamination, which in turn decreased public and private agreement. With the full sample, there was a direct effect of pressure predicting increased public agreement, but indirect effects found that pressure indirectly predicted decreased public agreement. With the non-posting only sample, even though there was no direct effect of pressure predicting public agreement, indirect effects reveal a backfiring effect where pressure indirectly predicted both decreased public and private agreement. Although the results show inconsistent results for direct backfiring effects for pressure, these data clearly suggest pressure *does* consistently backfire *indirectly* beneath the surface.

The backfiring effects of forced consensus go beyond agreement towards the given issue. Reactance and informational contamination also indirectly influenced people's attitudes towards LGBT persons and their concern for political correctness.

Indirect Effect of Pressure on LGBT Attitudes and Concern for Political Correctness

Posting one's preferred pronouns – which aims to support LGBT rights as well as diversity and inclusion - is considered politically correct (McBride, 2017; Murray, 2019). But what happens when people are forced to appear politically correct and support LGBT

persons? Mediation analyses found that - regardless of whether people currently post pronouns on Twitter or not – forcing everyone to post preferred pronouns on Twitter increases both reactance and informational contamination. These increases in turn predicted decreased LGBT attitudes and increased concerns for political correctness.

What does this imply?

People experience reactance when forced to do something. If people are forced to support LGBT persons by posting pronouns, they experience reactance and want to act against the perceived pressure. When they are forced to support LGBT persons, they report supporting them less.

People also experience informational contamination when they perceive an artificial consensus. If people are forced to support LGBT persons by posting pronouns, they perceive the forced pronoun posting as a politically-loaded agenda where people can no longer tell who is posting by choice and who is posting just to stay on Twitter. The artificial consensus that aims to support LGBT persons backfires and led to decreased LGBT attitudes.

Consensus Does Not Matter

Given the backfiring effects of forced consensus from previous work (e.g. Conway & Schaller, 2005; Conway et al., 2009; Conway & Repke, 2019), one hypothesized way to reduce informational contamination is to have participants believe that there is a genuine consensus (Conway et al., 2021). If participants think that most Twitter users support posting pronouns, then they will likely not see Twitter's policy of forced pronoun posting as an artificial consensus or reflective of some political agenda.

Indeed, almost half of the participants in the current sample (112 out of 260) already post preferred pronouns on Twitter.

But that does not seem to be the case. Analysis on the full sample and non-posting only sample all point to null effects (both direct and indirect) of consensus, suggesting that consensus had no influence on public agreement, private agreement, LGBT attitudes, and PC concerns. It is also possible that participants did not encode the consensus manipulation well given that it is a hypothetical scenario. It is certainly more difficult to imagine everyone agreeing on a divisive political issue than imagining how one reacts to a hypothetical policy.

Limitations of the Present Work

Like all studies, this one has limitations. This study only examined one type of politically-loaded cue as a marker for forced political allegiance. Preferred pronouns is a recent manifestation of an impression management cue in the online world, but its relevance and political connotation might change within the next decade. As the contentiousness of political issues is contingent on its time period (Sullivan, 2020), it is unclear how preferred pronouns will pan out in the future.

Another limitation is that this study only examined one social media platform – Twitter. Focusing only on Twitter as a marker for how people behave on social media platforms is limited because each social media platform has different norms of impression management and attracts different demographics. While Twitter was chosen in this study because many users discuss political issues on a daily basis, the Pew Research Center found that Twitter users in the U.S. are younger, more educated, and more left-leaning than the U.S. general public. The majority of Twitter users do not

tweet, and only around 10% who do tweet focus on politics (Wojcik & Hughes, 2019). As such, Twitter users' behavior cannot generalize to all U.S. adults.

Even though all of my sample's participants are Twitter users, asking how they react to a hypothetical policy is not the same as the reality of that policy. That being said, self-reports of hypothetical scenarios and questionnaires to gauge imagined reactions are becoming commonplace in the field of social psychology. As Baumeister et al. noted as far back as 14 years ago (2007), there has been an eclipse of studying actual observable behavior in social psychology.

Finally, the characteristics of this sample's pronoun-posters are not in line with theoretical expectations drawn from political psychology. Counterintuitively, those who posted pronouns self-reported to be more conservative and reported that posting pronouns make themselves appear conservative. It is puzzling why the Twitter pronoun-posters in this sample reported this way, which may call into question the replicability of this sample (assuming the sample is drawn from the same population) and by extension the generalizability of this sample's pronoun-posters.

Future Directions

As mentioned above, this study revealed backfiring effects of forced political allegiance on only one social media platform - Twitter. Future work could examine online platforms that focus on other aspects of one's life, such as work or relationships. Since each type of platform encourages users to highlight different aspects of the self, future research could examine how users of these different platforms react to political cues.

Forced political allegiance in the realm of work likely elicits strong backfiring effects. If people are forced to appear political on their company website, it might backfire more heavily than the Twitter pronoun scenario because the consequences of refusing to follow policies pertaining to career impression management could be much more pernicious (e.g. decreased chance of promotion to losing one's job) than getting suspended on social media. People can afford to lose access to one social media platform, but cannot as easily afford to ignore demands from one's workplace.

In contrast, forced political allegiance in the realm of online dating should elicit milder backfiring effects because the consequences of compromising one's first impression on a dating app (where people talk to strangers on a casual one-to-one basis) is not as serious as a conflict at work or on social media (where interacting with multiple groups simultaneously due to context collapse affects one's public reputation). That being said, there will still be some backfiring effects if people are forced to appear political on dating apps. Indeed, in the light of the divisive protests in Hong Kong, many Tinder users in Hong Kong have since voluntarily posted a "yellow ribbon" or "blue ribbon" to signal pro-democracy or pro-police respectively. While people who display this form of political allegiance aim to match with others who share a similar political view (Zheng, 2019), issues will arise if people are forced to post a particular ribbon online, as the ribbon serves as a cue for political ideology.

Additionally, future research could examine how influential leaders' choice to post preferred pronouns online has potential ripple effects. As influential leaders such as Tim Cook (CEO of Apple) have begun to post preferred pronouns on his Twitter bio in

December 2020 (Meisenzahl, 2020), it may shift norms or create implicit expectations for people within the same industry to follow. As more leaders and executives post pronouns, it may create a form of consensus or social proof that certain industries are in favor of posting pronouns.

Perhaps most important is tracking the effects of forced political allegiance in the ever-evolving sociopolitical context. The data for this study was collected and analyzed in January 2021, where the United States is reported to experience political sectarianism, meaning political outgroup hate is reported to be stronger than political ingroup love (Finkel et al., 2020). My data suggests that forced display of preferred pronouns on Twitter indirectly decreased LGBT attitudes and increased PC concerns, but these backfiring effects might be exacerbated if polarization continues, or attenuated if political parties begin to unite.

Concluding Remarks

The results from the current study suggests that pressuring people to post politically-loaded cues on Twitter could potentially backfire. On the surface, there will likely be drops in private agreement for those who do not currently adopt the politically-loaded cue. But on a broader level, forcing people to appear political will likely backfire because people may report negative attitudes towards the group whom the cue aims to support. Preferred pronouns aims to support LGBT individuals, but forced posting of preferred pronouns on social media - conceptualized in a hypothetical scenario - indirectly decreased LGBT attitudes and increased concerns for political correctness. There may be good reason to make people appear a certain way on Twitter, but

controlling one's online impression management comes with the cost of reactance and informational contamination.

References

- ANES. (2020, July 29). ANES Announcement: 2020 Pilot Study Data Available - ANES: American National Election Studies. Retrieved from <https://electionstudies.org/anes-announcement-2020-pilot-study-data-available/>
- Ashokkumar, A., Talaiifar, S., Fraser, W. T., Landabur, R., Buhrmester, M., Gómez, Á., ... & Swann Jr, W. B. (2020). Censoring political opposition online: Who does it and why. *Journal of Experimental Social Psychology, 91*, 104031.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science, 26*(10), 1531–1542.
- Baron, D. E. (2020). What's your pronoun?: Beyond he and she. New York ; London: Liveright Publishing Corporation.
- Barnidge, M., Gunther, A. C., Kim, J., Hong, Y., Perryman, M., Tay, S. K., & Knisely, S. (2020). Politically motivated selective exposure and perceived media bias. *Communication Research, 47*(1), 82–103.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on Psychological Science, 2*(4), 396-403.
- Bloom, P. (2017). *Against empathy: The case for rational compassion*. Random House.
- Brammer, J. P. (2019, July 22). 'She/hers': In progressive move, 3 presidential hopefuls add pronouns to their bios. Retrieved from <https://www.nbcnews.com/feature/nbc-out/she-hers-progressive-move-3-presidential-hopefuls-add-pronouns-their-n1032101>

- Chan, A., Madsen, R., & Unger, J. (2009). *Chen village: Revolution to globalization*. Univ of California Press.
- Chan, L., & Conway, L. G., III. (2018). Autocratic government moderates the relationship between culture and legal restriction. *Journal of Cross-Cultural Psychology, 49*, 1457-1463.
- Chan, L., McFarland, J. D., & Conway, L. G., III. (2018). Political contamination of social psychology: A review of Crawford and Jussim's (2017) edited book 'The politics of social psychology.' *Social Justice Research, 31*, 323-333.
- Charlesworth, T. E., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: Long-term change and stability from 2007 to 2016. *Psychological Science, 30*(2), 174-192.
- Cheng, C., Chan, L., & Chan, H.C.Y (2021). The social impact of 3G, 4G, and 5G on Hong Kong from 2006- 2020. In *Official Guide to ICT Industry in Hong Kong 2020* (pp. 31-75). Communications Association of Hong Kong: Hong Kong.
- Clark, C. J., Liu, B. S., Winegard, B. M., & Ditto, P. H. (2019). Tribalism is human nature. *Current Directions in Psychological Science, 28*(6), 587-592.
- Clayton, T. (1998). Building the new Cambodia: Educational destruction and construction under the Khmer Rouge, 1975–1979. *History of Education Quarterly, 38*(1), 1-16.
- Conway, L. G., III, Chan, L., & Woodard, S. R. (2019). Socio-ecological influences on political ideology. *Current Opinion in Psychology, 32*, 76-80.
- Conway, L. G, III, Houck, S. C., Chan, L., Repke, M. A., & McFarland, J. (2021). The agreement paradox: How pressures for agreement can ultimately divide us. In J.-

W. van Prooijen (Ed.), *Current Issues in Social Psychology: Political Polarization*. New York: Routledge.

- Conway, L. G., III, & McFarland, J. D. (2019). Do right-wing and left-wing authoritarianism predict election outcomes?: Support for Obama and Trump across two United States presidential elections. *Personality and Individual Differences, 138*, 84-87.
- Conway III, L. G., & Repke, M. A. (2019). The psychological contamination of pro-environmental consensus: Political pressure for environmental belief agreement undermines its long-term power. *Journal of Environmental Psychology, 62*, 12-21.
- Conway, L. G., Repke, M. A., & Houck, S. C. (2017). Donald Trump as a cultural revolt against perceived communication restriction: Priming political correctness norms causes more Trump support. *Journal of Social and Political Psychology, 5*(1), 244-259.
- Conway, L. G. III, Salcido, A., Gornick, L. J., Bongard, K. A., Moran, M., & Burfiend, C. (2009). When self-censorship norms backfire: The manufacturing of positive communication and its ironic consequences for the perceptions of groups. *Basic and Applied Social Psychology, 31*, 335-347.
- Conway, L. G., & Schaller, M. (2005). When authorities commands backfire: Attributions about consensus and effects on deviant decision making. *Journal of Personality and Social Psychology, 89*(3), 311-326.

- Conway, L. G., III, & Woodard, S. R. (2019). Integrative complexity across domains and across time: Evidence from political and health domains. *Personality and Individual Differences*. DOI: <https://doi.org/10.1016/j.paid.2019.109713>
- Conway III, L. G., Woodard, S. R., Zubrod, A., & Chan, L. (manuscript under review). *Why are conservatives less concerned about the coronavirus (COVID-19) than liberals? Testing experiential versus political explanations*. [Pre-print available at <https://psyarxiv.com/fgb84/>].
- Crawford, J. T. (2017). Are conservatives more sensitive to threat than liberals? It depends on how we define threat and conservatism. *Social cognition*, 35(4), 354-373.
- Crawford, J. T., & Jussim, L. J. (2018). *The politics of social psychology*. New York: Routledge, Taylor & Francis Group.
- Daniller, A. (2019, November 12). Americans' immigration policy priorities. Retrieved from <https://www.pewresearch.org/fact-tank/2019/11/12/americans-immigration-policy-priorities-divisions-between-and-within-the-two-parties/>
- Dembroff, R., & Wodak, D. (2018, June 4). If someone wants to be called 'they' and not 'he' or 'she', why say no? | Robin Dembroff and Daniel Wodak. Retrieved from <https://www.theguardian.com/commentisfree/2018/jun/04/gender-neutral-pronouns-they-he-she-why-deny>
- Duckitt, J., & Sibley, C. G. (2016). Personality, ideological attitudes, and group identity as predictors of political behavior in majority and minority ethnic groups. *Political Psychology*, 37(1), 109-124.

- Duggan, M., & Smith, A. (2016, October 25). Americans, Politics and Social Media. Retrieved from <https://www.pewresearch.org/internet/2016/10/25/the-political-environment-on-social-media/>
- Farrow, R. (2019). *Catch and kill: Lies, spies, and a conspiracy to protect predators*. Little, Brown, and Compant.
- Fingerhut, H. (2016, July 20). In 'political correctness' debate, most Americans think too many people are easily offended. Retrieved from <https://www.pewresearch.org/fact-tank/2016/07/20/in-political-correctness-debate-most-americans-think-too-many-people-are-easily-offended/>
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., ... & Druckman, J. N. (2020). Political sectarianism in America. *Science*, 370(6516), 533-536.
- Fractured Nation: Widening Partisan Polarization and Key Issues in 2020 Presidential Elections. (2019, October 20). Retrieved from <https://www.prii.org/research/fractured-nation-widening-partisan-polarization-and-key-issues-in-2020-presidential-elections/>
- Geiger, A. W., & Graf, N. (2019, September 5). About one-in-five U.S. adults know someone who goes by a gender-neutral pronoun. Retrieved from <https://www.pewresearch.org/fact-tank/2019/09/05/gender-neutral-pronouns/>
- Gentzkow, M., & Shapiro, J. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126, 1799–1839.
- Gil-Lopez, T., Shen, C., Benefield, G. A., Palomares, N. A., Kosinski, M., & Stillwell, D. (2018). One size fits all: Context collapse, self-presentation strategies and

- language styles on Facebook. *Journal of Computer-Mediated Communication*, 23(3), 127–145.
- Grubbs, J. B., Warmke, B., Tosi, J., James, A. S., & Campbell, W. K. (2019). Moral grandstanding in public discourse: Status-seeking motives as a potential explanatory mechanism in predicting conflict. *PloS one*, 14(10).
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.
- Haidt, J., & Rose-Stockwell, T. (2019, November 12). The Dark Psychology of Social Networks. Retrieved from <https://www.theatlantic.com/magazine/archive/2019/12/social-media-democracy/600763/>
- Harris, E. A., & Van Bavel, J. J. (2020, May 20). Preregistered Replication of “Feeling superior is a bipartisan issue: Extremity (not direction) of political views predicts perceived belief superiority”. <https://doi.org/10.31234/osf.io/hfuas>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis a regression- based approach*. New York, NY: Guilford Press.
- Heimlich, R. (2011a, January 10). Division Over Gun Control. Retrieved from <https://www.pewresearch.org/fact-tank/2011/01/10/division-over-gun-control/>
- Heimlich, R. (2011b, October 7). Same-Sex Marriage Issue Causes Divisions Among the Public and Within the Political Parties. Retrieved from <https://www.pewresearch.org/fact-tank/2011/10/07/same-sex-marriage-issue-causes-divisions-among-the-public-and-within-the-political-parties/>

- Houck, S. C., Conway, L. G., III, & Repke, M. (2014). Personal closeness and perceived torture efficacy: If torture will save someone I'm close to, then it must work. *Peace and Conflict: Journal of Peace Psychology*, 20, 590-592.
- Inbar, Y., & Lammers, J. (2012). Political diversity in social and personality psychology. *Perspectives on Psychological Science*, 7(5), 496-503.
- Jost, J. T., Nosek, B. A., & Gosling, S. D. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on Psychological Science*, 3(2), 126-136.
- Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020, February 21). U.S. Media Polarization and the 2020 Election: A Nation Divided. Retrieved from <https://www.journalism.org/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/>
- Kang, J., & Wei, L. (2020). Let me be at my funniest: Instagram users' motivations for using Finsta (aka, fake Instagram). *The Social Science Journal*, 57(1), 58-71.
- Kennedy, R., Clifford, S., Burleigh, T., Jewell, R., & Waggoner, P. (2018). The shape of and solutions to the MTurk quality crisis. *Available at SSRN 3272468*.
- Krämer, N. C., & Haferkamp, N. (2011). Online self-presentation: Balancing privacy concerns and impression construction on social networking sites. In *Privacy online* (pp. 127-141). Springer, Berlin, Heidelberg.
- Kuran, T. (1987). Preference falsification, policy continuity and collective conservatism. *The Economic Journal*, 97(387), 642-665.
- Landtsheer, C. L. D., De Vries, P., & Vertessen, D. (2008). Political impression management: How metaphors, sound bites, appearance effectiveness, and

- personality traits can win elections. *Journal of Political Marketing*, 7(3-4), 217-238.
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, 107(1), 34-47.
- Levendusky, M. (2013), Why do partisan media polarize viewers? *American Journal of Political Science* 57(3): 611–623.
- Logie, C., Bridge, T. J., & Bridge, P. D. (2007). Evaluating the phobias, attitudes, and cultural competence of master of social work students toward the LGBT populations. *Journal of Homosexuality*, 53(4), 201-221.
- Malloy, T., & Schwartz, D. (2020, March 9). Biden Crushes Sanders In Democratic Race, Quinnipiac University National Poll Finds; More Disapprove Of Trump's Response To Coronavirus. Retrieved from <https://poll.qu.edu/national/release-detail?ReleaseID=3657>
- Marwick, A. E., & Boyd, D. (2010). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133.
- McBride, J. (2017, November 21). A professor's refusal to use gender-neutral pronouns, and the vicious campus war that followed. Retrieved from <https://torontolife.com/city/u-t-professor-sparked-vicious-battle-gender-neutral-pronouns/>
- Meisenzahl, M. (2020, December 6). Apple CEO Tim Cook added his pronouns to his Twitter bio. Retrieved from <https://www.msn.com/en-us/money/other/apple-ceo-tim-cook-added-his-pronouns-to-his-twitter-bio/ar-BB1bENr1>

- Murray, D. (2019). *The Madness of Crowds: Gender, Race and Identity*. Bloomsbury Publishing USA.
- Paikin, Steve. "Genders, Rights and Freedom of Speech." *TVO*. The Agenda, 26 Oct. 2016. Web. 28 Nov. 2016. "<http://tvo.org/video/programs/the-agenda-with-steve-paikin/genders-rights-and-freedom-of-speech>"
- Parker, K., Horowitz, J. M., Igielnik, R., Oliphant, J. B., & Brown, A. (2019, December 31). Guns in America: Attitudes and experiences of Americans. Retrieved from <https://www.pewsocialtrends.org/2017/06/22/americas-complex-relationship-with-guns/>
- Pinker, S. (2018). *Enlightenment now: The case for reason, science, humanism, and progress*. Penguin.
- Pluckrose, H., & Lindsay, J. (2020). *Cynical theories: How universities made everything about race, gender, and identity-and why this harms everybody*. Great Britain: Swift Press.
- Popan, J. R., Coursey, L., Acosta, J., & Kenworthy, J. (2019). Testing the effects of incivility during internet political discussion on perceptions of rational argument and evaluations of a political outgroup. *Computers in Human Behavior*, 96, 123–132.
- Ranzini, G., & Hoek, E. (2017). To you who (I think) are listening: Imaginary audience and impression management on Facebook. *Computers in Human Behavior*, 75, 228–235.
- Rosling, H., Rönnlund, A. R., & Rosling, O. (2019). *Factfulness*. Paris: Flammarion.

- Saad, G. (2020). *The parasitic mind: How infectious ideas are killing common sense*. Washington, DC: Regnery Publishing.
- Schaeffer, K. (2020, March 04). Far more Americans see 'very strong' partisan conflicts now than in the last two presidential election years. Retrieved from <https://www.pewresearch.org/fact-tank/2020/03/04/far-more-americans-see-very-strong-partisan-conflicts-now-than-in-the-last-two-presidential-election-years/>
- Schneider, J. (2020, July 16). Trump finally wore a mask. Why do conservatives resist them? Retrieved from <https://www.washingtonpost.com/politics/2020/07/16/trump-finally-wore-mask-why-do-conservatives-resist-them/>
- Shore, M. (1996). The Sacred and the Myth: Havel's greengrocer and the transformation of ideology in communist Czechoslovakia. *Contagion: Journal of Violence, Mimesis, and Culture*, 3(1), 163-182.
- Shrier, A. (2020). *Irreversible damage: The transgender craze seducing our daughters*. Washington, DC: Regnery Publishing.
- Siu, P., & Lau, C. (2020, May 24). Tear gas fired, arrests made as thousands protest against national security law. Retrieved May 26, 2020, from <https://www.scmp.com/news/hong-kong/politics/article/3085822/hong-kong-opposition-activist-arrested-hundreds-gather>
- Soh, D. (2020). *The end of gender: Debunking the myths about sex and identity in our society*. New York: Threshold Editions.
- Sterling, J., Jost, J. T., & Bonneau, R. (2020). Political psycholinguistics: A comprehensive analysis of the language habits of liberal and conservative social

- media users. *Journal of Personality and Social Psychology*. Advance online publication.
- Sullivan, D. (2020). Social psychological theory as history: Outlining the critical-historical approach to theory. *Personality and Social Psychology Review*, 24(1), 78-99.
- Susskind, J. (2018). *Future politics: Living together in a world transformed by tech*. Oxford University Press.
- Swank, E., Woodford, M. R., & Lim, C. (2013). Antecedents of pro-LGBT advocacy among sexual minority and heterosexual college students. *Sexuality Research and Social Policy*, 10(4), 317-332.
- The Daily Wire. (2020, Nov. 9). *Student challenges Ben Shapiro on Transgenderism: "If it makes them happier, what's the harm?"* Retrieved from: <https://www.youtube.com/watch?v=VXAbDXdOLa8>
- Varava, K. A., & Quick, B. L. (2015). Adolescents and Movie Ratings: Is psychological reactance a theoretical explanation for the forbidden fruit effect? *Journal of Broadcasting & Electronic Media*, 59(1), 149-168.
- Walther, J. B., Heide, B. V. D., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on Facebook: Are we known by the company we keep? *Human Communication Research*, 34(1), 28-49.
- Wareham, J. (2019) Should you put pronouns in Email signatures and social media Bios? *Forbes*. Retrieved

from <https://www.forbes.com/sites/jamiewareham/2020/12/30/should-you-put-pronouns-in-email-signatures-and-social-media-bios/#2273ff5d6320>

Wojcik, S., & Hughes, A. (2019, April 24). How Twitter users compare to the general public. Retrieved from <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/>

Zheng, M. (2019, December 14). Protest dating shake-up: Yellow and blue don't match for Hong Kong's lovers. Retrieved from <https://www.scmp.com/news/hong-kong/politics/article/3041932/new-hong-kong-dating-game-i-you-are-you-yellow-or-blue>

Table 1

Mean Ratings by Pressure and Consensus for Full Sample (SD) (N = 260)

	Pressure		Consensus	
	No	Yes	No	Yes
Public	4.78	5.11	4.89	5.13
Agreement	(1.40)	(1.50)	(1.49)	(1.46)
Private	4.86	4.75	4.78	4.79
Agreement	(1.44)	(1.76)	(1.67)	(1.67)
LGBT	5.12	4.50	4.64	4.76
Attitudes	(1.56)	(1.48)	(1.49)	(1.58)
PC	2.90	3.49	3.39	3.22
Concerns	(1.32)	(1.16)	(1.23)	(1.25)

Table 2

Mean Ratings by Pressure and Consensus for Pronoun-Posters-Only Sample (SD)

(N = 112)

	Pressure		Consensus	
	No	Yes	No	Yes
Public	5.54	5.73	5.70	5.68
Agreement	(.73)	(.71)	(.60)	(.83)
Private	5.32	5.62	5.61	5.52
Agreement	(.92)	(.73)	(.76)	(.80)
LGBT	4.67	3.74	3.83	4.00
Attitudes	(1.36)	(.98)	(.93)	(1.28)
PC	3.19	3.89	3.91	3.60
Concerns	(1.25)	(.97)	(.79)	(1.27)

Table 3

Mean Ratings by Pressure and Consensus for Non-Posting Only Sample (SD) (N = 148)

	Pressure		Consensus	
	No	Yes	No	Yes
Public	4.52	4.46	4.26	4.71
Agreement	(1.48)	(1.82)	(1.66)	(1.68)
Private	4.70	3.84	4.14	4.25
Agreement	(1.56)	(2.04)	(1.89)	(1.92)
LGBT	5.28	5.31	5.27	5.32
Attitudes	(1.60)	(1.49)	(1.54)	(1.55)
PC	2.80	3.08	2.99	2.95
Concerns	(1.34)	(1.20)	(1.36)	(1.17)

Table 4

The impact of reactance: Simple and indirect effects of pressure, consensus, and posted pronouns on public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness (Full Sample)

	Indirect Effect Via Reactance			
	Simple Effect	Indirect Effect	Indirect Lower CI	Indirect Upper CI
Pressure/Public Agreement	.54*	-.21*	-.40	-.04
Pressure/Private Agreement	.01	-.12	-.32	.08
Pressure/LGBT Attitudes	.06	-.68*	-.94	-.45
Pressure/PC Concerns	-.09	.68*	.47	.91
Consensus/Public Agreement	.21	.03	-.01	.09
Consensus/Private Agreement	-.01	.02	-.02	.10
Consensus/ LGBT Attitudes	-.01	.13	-.07	.34
Consensus/ PC Concerns	-.03	-.13	-.34	.07
Posted Pronouns/Public Agreement	1.62*	-.41*	-.61	-.23
Posted Pronouns/Private Agreement	1.79*	-.42*	-.64	-.22
Posted Pronouns/ LGBT Attitudes	-.83*	-.55*	-.79	-.34
Posted Pronouns/ PC Concerns	.13	.67*	.47	.88

Note: $N = 260$ * $p < .05$; Confidence intervals based on 5000 bootstrapped samples.

Table 5

The impact of informational contamination: Simple and indirect effects of pressure, consensus, and posted pronouns on public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness (Full Sample)

	Indirect Effect Via Informational Contamination			
	Simple Effect	Indirect Effect	Indirect Lower CI	Indirect Upper CI
Pressure/Public Agreement	.47*	-.14*	-.30	-.01
Pressure/Private Agreement	-.06	-.05	-.21	.11
Pressure/LGBT Attitudes	-.14	-.48*	-.71	-.26
Pressure/PC Concerns	.10	.49*	.28	.70
Consensus/Public Agreement	.22	.02	-.02	.09
Consensus/Private Agreement	.03	.01	-.03	.07
Consensus/ LGBT Attitudes	.03	.09	-.11	.30
Consensus/ PC Concerns	-.08	-.09	-.29	.11
Posted Pronouns/Public Agreement	1.51*	-.30*	-.48	-.14
Posted Pronouns/Private Agreement	1.62*	-.26*	-.45	-.09
Posted Pronouns/ LGBT Attitudes	-.94*	-.44*	-.65	-.27
Posted Pronouns/ PC Concerns	.26*	.53*	.35	.72

Note: $N = 260$ * $p < .05$; Confidence intervals based on 5000 bootstrapped samples.

Table 6

The impact of reactance: Simple and indirect effects of pressure and consensus on public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness (Non-posting only Sample)

	Indirect Effect Via Reactance			
	Simple Effect	Indirect Effect	Indirect Lower CI	Indirect Upper CI
Pressure/Public Agreement	.46	-.52*	-.86	-.24
Pressure/Private Agreement	-.41*	-.46*	-.82	-.18
Pressure/LGBT Attitudes	.35	-.33*	-.59	-.13
Pressure/PC Concerns	-.21	.48*	.26	.73
Consensus/Public Agreement	.37	.08	-.15	.32
Consensus/Private Agreement	.02	.09	-.15	.35
Consensus/ LGBT Attitudes	.01	.05	-.08	.23
Consensus/ PC Concerns	.04	-.08	-.32	.14

Note: $N = 148$ * $p < .05$; Confidence intervals based on 5000 bootstrapped samples.

Table 7

The impact of informational contamination: Simple and indirect effects of pressure and consensus on public agreement, private agreement, attitudes towards LGBT persons, and concern for political correctness (Non-posting only Sample)

	Indirect Effect Via Informational Contamination			
	Simple Effect	Indirect Effect	Indirect Lower CI	Indirect Upper CI
Pressure/Public Agreement	.21	-.27*	-.54	-.05
Pressure/Private Agreement	-.65*	-.21*	-.45	-.03
Pressure/LGBT Attitudes	.24	-.22*	-.46	-.04
Pressure/PC Concerns	-.00	.28*	.06	.51
Consensus/Public Agreement	.48	-.03	-.24	.18
Consensus/Private Agreement	.13	-.02	-.23	.17
Consensus/ LGBT Attitudes	.08	-.02	-.18	.16
Consensus/ PC Concerns	-.07	.03	-.20	.25

Note: $N = 148$ * $p < .05$; Confidence intervals based on 5000 bootstrapped samples.

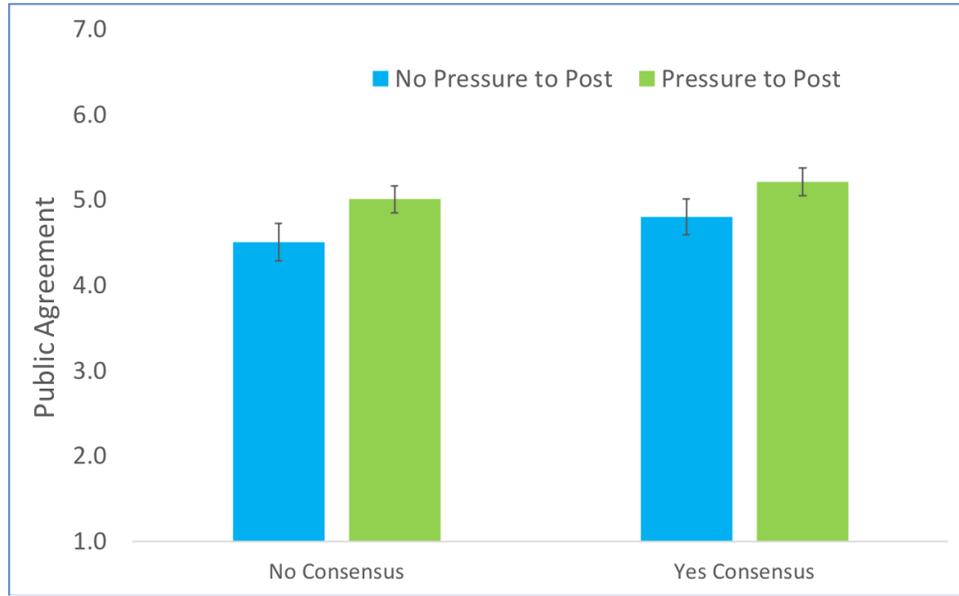


Figure 1. Effects of Pressure and Consensus on Public Agreement (Full Sample). Error bars represent standard error of the mean.

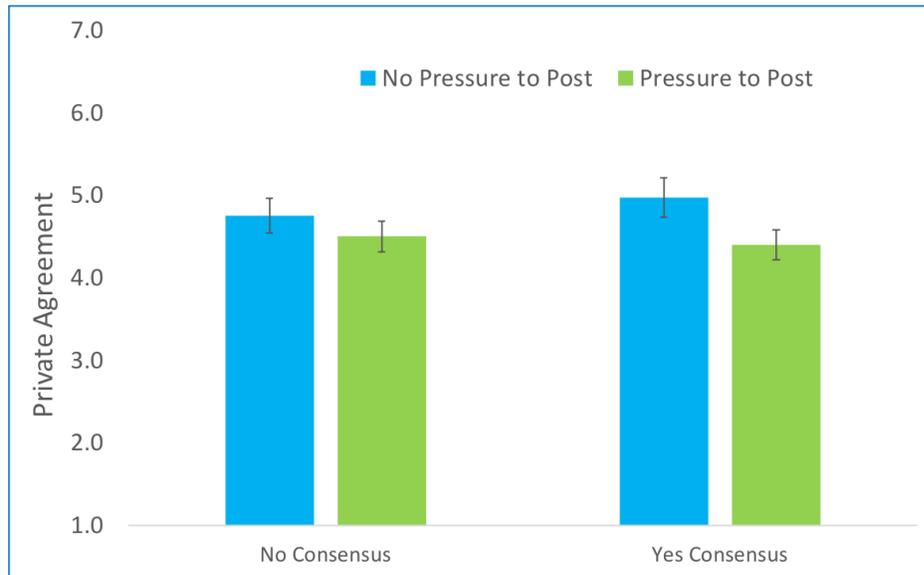


Figure 2. Effects of Pressure and Consensus on Private Agreement (Full Sample). Error bars represent standard error of the mean.

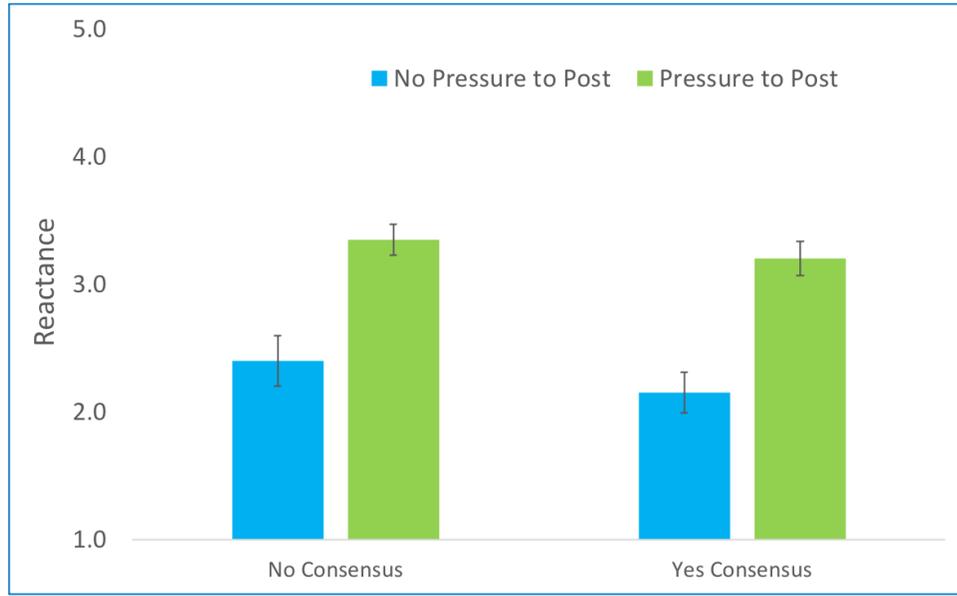


Figure 3. Effects of Pressure and Consensus on Reactance (Full Sample). Error bars represent standard error of the mean.

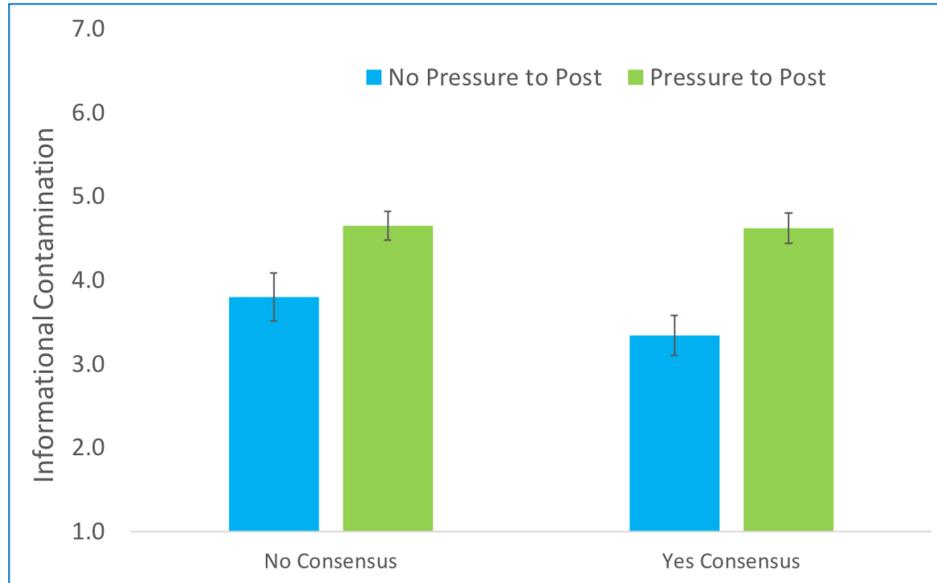


Figure 4. Effects of Pressure and Consensus on Informational Contamination (Full Sample). Error bars represent standard error of the mean.

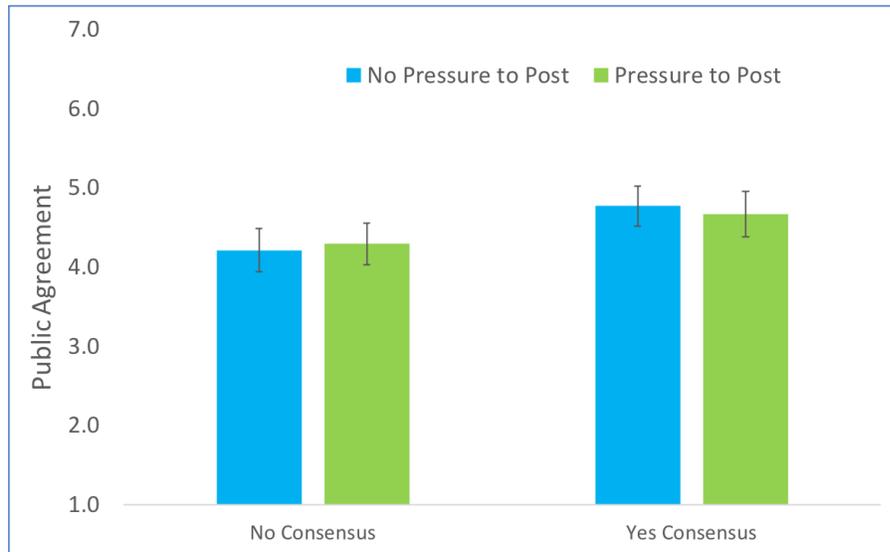


Figure 5. Effects of Pressure and Consensus on Public Agreement (Non Pronoun Posting Only Sample). Error bars represent standard error of the mean.

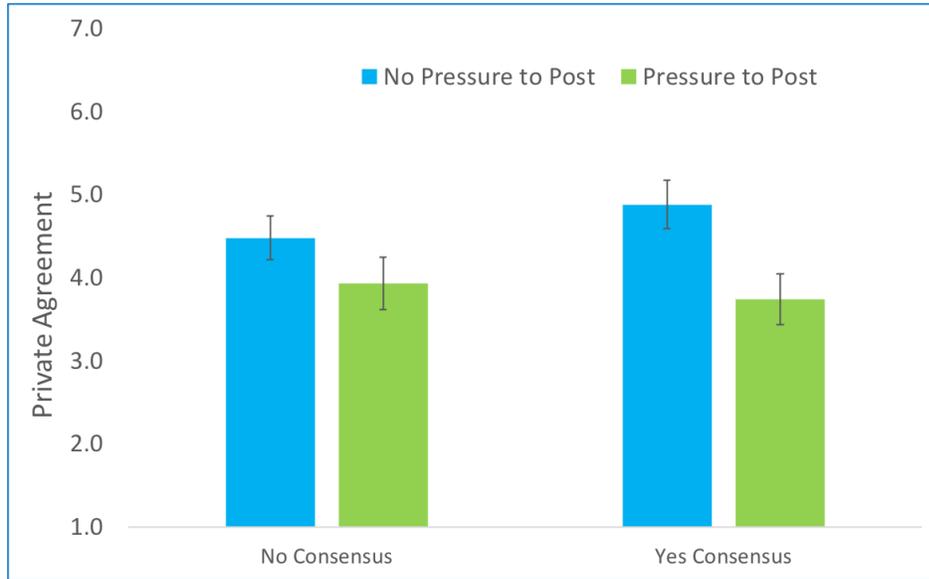


Figure 6. Effects of Pressure and Consensus on Private Agreement (Non Pronoun Posting Only Sample). Error bars represent standard error of the mean.

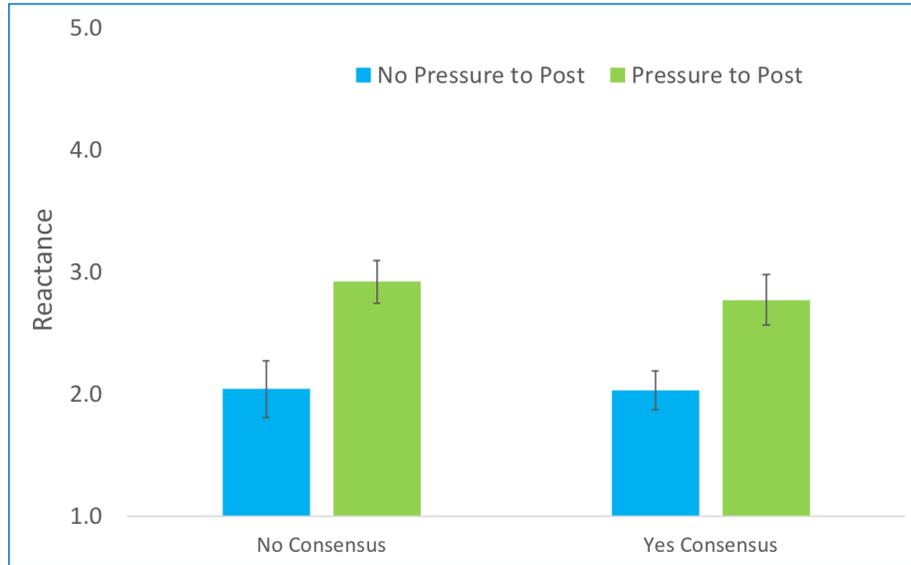


Figure 7. Effects of Pressure and Consensus on Reactance (Non Pronoun Posting Only Sample). Error bars represent standard error of the mean.

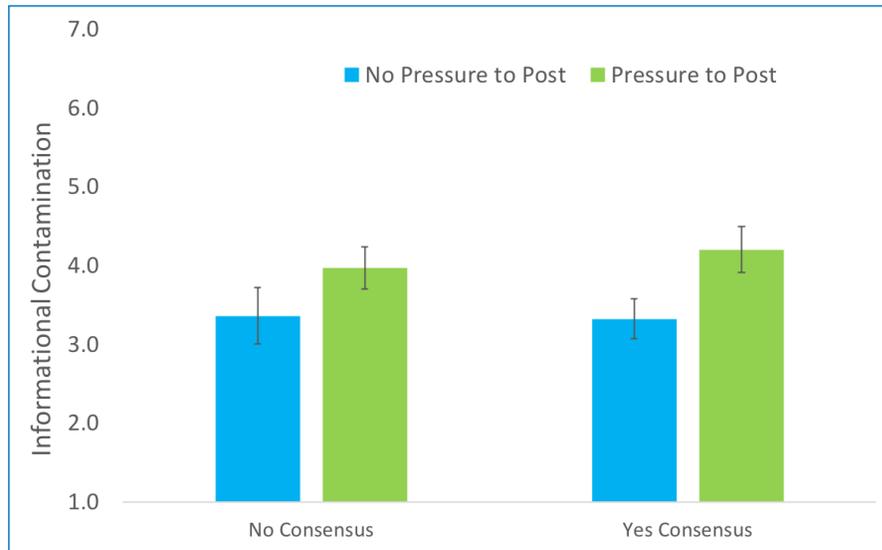


Figure 8. Effects of Pressure and Consensus on Informational Contamination (Non Pronoun Posting Only Sample). Error bars represent standard error of the mean.

Appendix A

Common Opening	Imagine that Twitter is implementing a new policy that aims to increase	
IVs	Yes	No
Pressure to Declare Pronouns (IV1)	<p>diversity and inclusion. Starting in a couple of months, this policy will require every Twitter user to declare one’s own preferred pronouns (e.g. he/him, she/her, they/them) on one’s Twitter bio or be suspended from using Twitter. There will be no exceptions.</p> <p>Now imagine that knowledge of this impending policy has already prompted many Twitter users to post preferred pronouns (e.g. he/him, she/her, they/them) on their profile.</p>	<p>user engagement. Starting in a couple of months, this policy will allow every Twitter user to customize their user interface with any Twitter-created or user-created background theme (e.g. sunsets, cute animals, city landscapes, superheroes) at no cost at all. This is completely optional and is up to the user to use (or not use).</p> <p>Although the policy is about background themes, some Twitter users have voluntarily chosen to post their preferred pronouns (e.g. he/him, she/her, they/them) on their profile.</p>
Consensus About Declaring Pronouns (IV2)	A recent poll that surveyed thousands of Twitter users found that although more liberals than conservatives support posting one’s preferred pronouns, the Twitter community generally supports posting preferred pronouns.	Twitter has not conducted a survey on who supports or opposes posting preferred pronouns on Twitter profiles, so we do not know how the Twitter community feels about preferred pronouns.

Appendix B

Attitude towards LGBT Persons (Logie et al., 2007)

Participants respond on a 5-point scale: 1 = strongly disagree, 5 = strongly agree.

1. If a person has homosexual feelings, they should do everything to overcome these feelings. (Reverse-scored)
2. Bisexuality is merely a different kind of lifestyle that should not be condemned.
3. Homosexuality is merely a different kind of lifestyle that should not be condemned.
4. Bisexuality is a threat to many of our basic social institutions. (Reverse-scored)
5. If a person feels that they belong to a different gender than the one they were born into, they should do everything to overcome these feelings. (Reverse-scored)
6. Transgender people threaten many of our basic social institutions. (Reverse-scored)
7. If a person has bisexual feelings, they should do everything to overcome these feelings. (Reverse-scored)
8. Homosexuality is a threat to many of our basic social institutions. (Reverse-scored)
9. Transgender people merely have a different sexual identity that should not be condemned.