UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

UNMASKING THE EFFECTS OF SOCIAL IDENTITY, POLARIZATION, AND ANGER ON INFORMATION PROCESSING AND ATTITUDES: AN EXPLORATION INTO THE 'MASK DEBATE' OF COVID-19 ON TWITTER

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

KEITH STRASBAUGH

Norman, Oklahoma

2022

UNMASKING THE EFFECTS OF SOCIAL IDENTITY, POLARIZATION, AND ANGER ON INFORMATION PROCESSING AND ATTITUDES: AN EXPLORATION INTO THE 'MASK DEBATE' OF COVID-19 ON TWITTER

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Shane Connelly, Chair

Dr. Jorge Mendoza

Dr. Claude Miller

Dr. Michael Mumford

Acknowledgements

The last ten years of undergraduate and graduate education at the University of Oklahoma have been transformative in many ways and, while numerous people have positively shaped my journey and this final educational endeavor at OU, there are several who need to be acknowledged. First, my graduate advisor, Dr. Shane Connelly, had a pivotal impact on my development, both personally and professionally. Through her infinite kindness, fierce belief in her students, and unending support, I was able to beat many odds that worked against me and achieve this long-term goal and ambition that reflects the work of generations of my family. Next, Drs. Jennifer Barnes, Jorge Mendoza, Claude Miller, and Michael Mumford, all of whom served on my graduate committees throughout my years at OU and provided both sage advice and much wisdom that brought me to where I am today. Without their belief in me, I would not have been able to pursue my dreams. Dr. Jennifer Wright also played a pivotal role in not only my academic and professional development, but also in the trajectory of my overall career aspirations. It is through her guidance and mentorship that I found the field of Customer Experience (CX), and her continuous encouragement facilitated the pursuit of the final steps towards my degree. Finally, Dr. Tobi Thompson, my eleventh-grade English teacher and friend, who both saved my life and never stopped believing in me is important to acknowledge and thank as well. She taught me many valuable life lessons that continue to help me today and guided me through the process of getting into college and managing and overcoming the many challenges that stood in my way.

This dissertation is dedicated not only to this enduring group of educators and mentors, but also to my friends and family throughout undergrad and grad school as well. Hayley Campbell & Skyler Flewellen, Elizabeth Monzingo, Amy Monzingo, Lynze Jones, Linzee &

Jeremy Sweet, Shawntal Brown, Marina Mery, Michelle Todd, Alise Dabdoub, Cassidy Krantz, Jessica Franklin, Ashley Jorgensen, Chanda Sanders, Madhuri Ramasubramanian, Divya Patel, Nosi Khumalo, Paxton Holt, Will Spain, Breea Clark, Brittany Mayes, Lexi Walton, Amy & Angelos Lambis, Bryton & Elriné Thompson, and Christen & Grace Raines, among many more cherished friends and loved ones, taught me how to pursue my dreams and never gave up on me. This extraordinary group of people I have had the honor of receiving support from throughout my journey are the reason I am able to write this dedication. And finally, to my partner, Shane Brown, with whom I share a family of rambunctious cats, numerous adventures, and a life that is enriching, motivating, and inspiring, this dissertation is dedicated to you.

Thank you, all of you, for supporting me in the pursuit of my Ph.D. In Industrial and Organizational Psychology.

Table of Contents

Abstract

The ‘mask debate’ surrounding COVID-19 (i.e., whether or not to wear a mask) has generated contention in online forums such as Twitter, where observed division of perspectives often runs along lines of social identity (e.g., political affiliation). Much of the research prior to COVID-19 suggests that social identity and the extent to which social media posts are polarizing exerts differential effects on user experiences. However, little research examines how social identity affects information processing of online COVID-19 discussions. Additionally, the question of how anger reactions, an emotion commonly explored in Twitter studies, may function on a more specific level (i.e., righteous vs. self-righteous) remains underexamined as well. Participants read and responded to a simulated Twitter feed focusing on the COVID-19 mask debate. Content polarization and explicit references to social identity groups were manipulated to assess their effects on three outcomes: 1) righteous and self-righteous anger, 2) biased processing of information contained in the Twitter feed (i.e., confirmation bias, selective attention bias, emotional content bias, commitment bias), and 3) attitudinal stance and strength regarding the topic of the Twitter feed. Additionally, righteous and self-righteous anger were explored as mediators of the relationships of social identity and polarization to cognitive biases and attitudes about the topic. Findings indicate effects of social identity on self-righteous anger, but not righteous anger, and polarization on strength of attitudes, confirmation bias and emotional content bias. Implications and future directions are discussed.

Keywords: Social identity; polarization; anger; cognitive biases

Unmasking the Effects of Social Identity, Polarization, and Anger on Information Processing and Attitudes: An Exploration into the 'Mask Debate' of COVID-19 on Twitter

Efforts to prevent the spread of the coronavirus (COVID-19) by wearing masks has generated much debate in online settings such as Twitter (Bhasin et al., 2020), in which social media users are heavily divided regarding when, what kind, and how to properly wear a mask (e.g., Al-Ramahi et al., 2021). These discussions are contentious because they are fueled by different values and beliefs about what constitutes the right thing to do (e.g., preserve individual freedoms vs. protect societal health). Over time, polarization and volume of tweets (i.e., limited character, posted content on Twitter) have significantly increased, with politicization among party lines of conservatives and liberals, Democrats and Republicans, driving much of the contentious discourse (Sanders et al., 2020). As the conversation progresses into questions of how to address communication of effective methods for mitigating the spread of COVID-19 around the world (e.g., vaccines, social distancing, wearing masks), it is important to understand the psychological processes underlying the division of viewpoints and how these influence affective, cognitive, and behavioral outcomes.

One such process is proposed by Roy and Ghosh (2021), who found through an analysis of 400,000 global tweets that different aspects of one's social identity may influence receptivity to intervention techniques to reduce COVID-19 cases, with blue states (i.e., Democratic majority) showing more acceptance of these preventive efforts in online discussions than red states (i.e., Republican majority). As a result, the type of content one engages with (e.g., how polarizing content is) and social identity may play important roles in how one responds to social media content in Twitter. These responses may be a function of emotional and cognitive responses to viewing polarized and social identity information in Twitter, but this has yet to be

tested within a causal framework. In social media spaces, it is common for users' social identities to be activated and for them to seek information that echoes and supports their perspectives and that may be contentious with other users (Karlsen et al., 2017). As a result, the original post by a user and their response sets the tone for many online discussions, and often produces effects on how subsequent responses to their post(s) occur in terms of emotions felt and expressed by other users (e.g., Li & Xiao, 2020). Emotions (i.e., anger) may act as mediators in this relationship of post content (i.e., social identity and polarization) and subsequent responses, such as attitude towards the issue and biases of information processed relative to content congruent with one's view. The effects of content that drives how a user responds to an original post or feed may be strengthened or weakened depending on the emotion (i.e., righteous vs. self-righteous anger) induced—a relationship that is underexplored in the literature.

This study makes several theoretical and practical contributions to the literature on psychological processes and social media. With respect to theory, this research adds to the understanding of the influences of the effects of social identity and polarization on emotional responses (i.e., anger) and biased information processing (e.g., confirmation bias, selective attention bias) in a social media context. Regarding emotion as an outcome, much of the research has broadly examined anger as a product of contention in online discussions; however, the extent to which different types of anger (i.e., righteous vs. self-righteous) are induced by polarization and social identity is less known in these settings. As a result, this research contributes to the emotion literature by establishing and exploring different types of anger as an outcome and as mediators of relationships of social media content and information processing to attitudes and responses in social media. Regarding information processing, this study examines the joint effects of polarization and social identity on confirmation bias, selective attention bias, emotional

content bias, and commitment bias, respectively. Accordingly, this research contributes to literature on social media engagement and contention/online debate by highlighting how certain content features of social media exchange shape one's ability to effectively process information (e.g., less biased, more accurate interpretations of arguments) and engage in discourse with others. This work also may have practical implications for structuring conversations in social media about contentious topics such as COVID-19. Results of this study could potentially be applied to improve communication with others in these spaces and, through increasing awareness of potential biases that can lead to misinterpreting or failing to understand a discussion in its entirety when engaging in online discourse. Additionally, through understanding users and how they respond to certain types of content, we may be better able to shape communication around preventative measures for COVID-19 by how these incorporate or reflect social identities.

Social Identity in Tweets and User Experience

Online discussions yield rich information about social media users, such as aspects of their systems of beliefs and attitudes towards certain topics. Bennett (2012) explores online political discourse through analyses of communication trends and proposes that these environments are unique avenues for users to share their opinions on political hot-topics, including environmentalism, fair trade, and social inequality, and to personalize their approach to expressing aspects of their identity. Yet, in these identity expressions, additional research highlights the need people have to connect with others on communal levels or to express social identity—with users of Twitter sometimes infusing their communication styles or posts with references to social groups or audiences with whom they identify as similar when writing posts (Tamburrini, Cinnirella, Jansen, & Bryden, 2015). These are self-expressions that simultaneously seek connection with other individuals who hold one or more similar social identities. In this

balance of self and group identification, respectively, users actively seek likeminded individuals, has the potential to create a sort of echo chamber that strengthens or supports a user's perspective(s), while ignoring or disregarding the opinions of those who may not agree with their ideas or beliefs (Garimella, De Francisci Morales, Gionis, & Mathioudakis, 2018). These relationships highlight the lack of balance users often experience when participating in (political) discourse on social media. Yet, how one identifies in these discussions and, subsequently, engages with others is a complex process that involves different facets of self-identification. The research on self-identification and engagement with others is vast and spans nearly forty years of extensive theory and experimentation (e.g., Hornsey, 2008; Stets & Burke, 2000; Trepte & Loy, 2006). One of the core concepts to emerge in this literature is Social Identity Theory (SIT).

SIT posits interactions with others exist on a spectrum ranging from highly interpersonal (e.g., self-conceptions focus on the individual with little reference to social categories) or highly intergroup (e.g., self-conceptions rely on categorizations into one or more social groups with little individualization) and that going from one end to the other shifts how we see ourselves and others (Hornsey 2008; Tajfel & Turner, 1979). Additionally, SIT proposes that we favor those we see as similar to ourselves—the 'us' relative to 'them' when conceptualizing identification with a group—and that the more we identify with a group (i.e., the more we move on the spectrum towards 'intergroup'), the more salient differences will become to us (Hornsey, 2008). During times of uncertainty, such as what many are experiencing during COVID-19, self-categorization into social groups may help people reduce uncertainty and build feelings of self-worth (Hogg, 2004). When social identities are activated, uncertainty is reduced because social groups offer prescriptions for beliefs, attitudes, feelings, and behavior. This research speaks to

the extent in which the strength of social identity may potentially impact the way in which one interacts with others who are either part of or not part of the in-group in online settings.

Examples of SIT in social media may be observed in online political debates, in which issues and social identification critically differentiate and divide people with different political group identities. For example, Garimella et al. (2017) found that Twitter discourse structure and content vary along strength of social identity. The 2016 U.S. presidential election is also an example of in-group vs. out-group differentiation and threats to social identity, in which political discussion became salient and turbulent and we observed more debate-centric discourse than previous elections on social media (e.g., Lau, Bligh, & Kohles, 2020). Much of this conversation centered around 'us' vs 'them' mentality, or how strongly social identity became salient, and divisions along political party affiliation lines (e.g., Republican vs. Democrat). For instance, if one identified as a Democrat and primarily interacted with other Democrats, this strengthened self-identification and favoring their party's presidential candidate, but still maintained an element of self-identification or individual identity (e.g., perceptions of gender differences, attitudes towards certain types of leadership preferences) (Lau, Bligh, & Kohles, 2020). However, if they interacted with Republicans or other political parties contrasting to their viewpoints on a contentious issue, their identification as a Democrat increased while aspects of their individual identity weakened because greater threat to social identity was perceived (Lau, Bligh, & Kohles, 2020).

Social Identity Theory and processes are useful for understanding behavior in social media settings such as social interactions (i.e., posts), and agreement/disagreement with opinions and information posted by other social media users, and the strength of these relationships, the actual content of the online posts that generate discourse are important to consider as well.

Specifically, the extent to which this content is polarizing may increase in-group identification, cohesion, and outgroup division, influencing interaction with other online forum users.

Polarization of Twitter Post Content and User Experience

The content people view online can act as a prime for processes of social identity, and the extent to which this content is either highly congruent or highly incongruent with group identity may lead to effects regarding online social media interaction outcomes as well. This content polarization may drive greater ingroup identification and depersonalization, in which the more evident differing opinions on issues become (i.e., 'us' vs 'them' mentality), the more likely one may be to reinforce beliefs and interact with members of the same group, while disregarding or denigrating another group's commentary (e.g., Mackie, 1986). That is, the more a Twitter feed presents contentious, polarized perspectives, the greater the likelihood that members of an online debate will 'select a side' that strengthens their perspective. However, it is important to note that this does not always occur. In some cases, political polarization on Twitter can also allow users the opportunity to grow their mindset and, in some situations, sway perspectives when political ideologies are not vastly dissimilar and communication is more agreeable and less defensive (e.g., social identity differences are not as salient in the discussion) (Gruzd, 2012). Online platforms have the utility of connecting users regardless of physical distance; yet it is important to highlight that a potential (common) outcome is the creation of 'perspective echo chambers' (Hong & Kim, 2016). This trend often occurs with individuals who are highly active in political discourse online and frequently interject opposing opinions on others'commentary thread or in response to others' social media posts (Conover, Ratkiewicz, Francisco, Goncalves, Flammini, & Menczer, 2011). This, in turn, may either result in numerous homogenous-perspective users voicing agreement or opposing-perspective users voicing disagreement.

An important component of the content that drives this contention is the extent that a social media poster expresses an opinion that is polarizing in tandem with the strength of their social identities expressed. This interaction between polarization and social identity may produce effects on cognition (i.e., anger, cognitive biases) that may impact reactions, abilities to accurately interpret and understand the discussions at hand, and even shape potential responses to these social media feeds.

Social Identity Twitter Post Content and User Emotion

The research findings regarding the effects of engaging with online political debate on emotions reveals several gaps in the literature and areas in need of further study. Chmiel and colleagues (2011) initially examined the effects of emotions expressed online through discussions and found that when a user expressed positive or negative emotion in their post, it often affects how others interact with that user online, resulting in emotional contagion. Expanding upon this idea of emotion contagion in online debates, Li and Xiao (2020) examined the prevalence of discrete emotions in these types of online interactions and found that the topic discussed affects the specific type of emotion displayed, whether negative (i.e., anger, disgust, fear, sadness), positive (i.e., joy), or neutral, and that immediate comments to a user mirrored the emotion in the original post. Wollebaek and colleagues (2019) proposed that those who become angry will seek information that confirms their views and will engage with posts that either support or contradict their perspectives; whereas users who experience anxiety as a product of a post will often narrow their focus to information that contradicts their perspective(s). These findings highlight a trend in the literature, in which much of the research regarding online discourse often focuses on negative emotions such as anger in a broad sense (e.g., Chen & Ng, 2017; Gregory, 2018), rather than examining more specifically how different types of anger (i.e.,

righteous anger vs. self-righteous anger) may initially be products of these online social interactions that exert impacts on social media interactions. While research proposes that anger in general may exert strong effects on cognitive outcomes (e.g., Wollebaek et al., 2019), it is unclear if more specific types of anger may produce similar or dissimilar effects to the same degree of anger in general.

The extent to which online discussions may produce certain emotions depends, in part, on the type of content a user interacts with on forums such as Twitter (Li & Xiao, 2020). Often, in the expression of social identity, it is common for users to become defensive and to express anger when beliefs and values related to that aspect of their identity feel challenged. (Wollebaek et al., 2019). However, anger is a complex emotion. Anger can be righteous, driven by a desire to see justice for victims of some wrongdoing or it can be self-righteous, driven by defensive/self-protective motives aimed at preserving one's moral character (Rothschild & Keefer, 2018). Righteous anger can be conceptualized as an extension of a negative emotion that motivates a need to correct and/or prevent a perceived injustice that has or will occur against society or a group of people (Tripp & Bies, 2009), whereas self-righteous anger is more focused on needing to correct and/or prevent a perceived injustice unique to one's in-group and social identity (Rothschild & Keefer, 2018). In a social media context, righteous anger may be observed in the 'mask debate' when a user(s) responds to a perspective that may pose a threat to society (e.g., not wearing a mask may endanger us all; wearing a mask is an infringement on people's freedom in general), whereas self-righteous anger is more focused on preserving and protecting one's in-group perspective to maintain one's own self-worth. Righteous anger can act as a boost of one's self-perceptions by restoring feelings of justice after an injustice has occurred (Green et al. 2019); whereas self-righteous anger focuses on alleviating feelings of personal guilt about the

role one's in-group group has played in a societal injustice. This anger is often directed at outgroups to remove blame that could potentially be directed at one's own in-group and thus oneself (Rothschild & Keefer, 2018). When social identity is present or highlighted in contentious social media exchanges, self-righteous anger may be more likely to occur than when social identity or in-group information is not included. Alternatively, righteous or justice driven anger may be more likely when social identity is not explicitly mentioned or highlighted (Nigmatullina & Bodrunova, 2018). For example, if one is exposed to a discussion of whether to wear a mask in light of COVID-19, self-righteous anger may result if the discussion seems to 'attack' one's social identity (i.e., Democrat vs. Republican). Branscombe and colleagues (1999) explain that social identity threat occurs when one's distinct social identity is challenged by another outgroup or outgroup member through perceptions of an insult, a lack of freewill, dismissive behavior towards their perspectives, or feeling inadequate as a product of how the other group is challenging them. The experience of these threats to social identity in online discussions may produce certain emotional reactions that will vary based upon how strongly one identifies with a particular group(s). Similarly, Chen and Ng (2017) found that when comments from others in political debates are upsetting or insulting towards an individual, anger is often induced.

However, the strength of identification with a particular social group or in-group is also likely to affect the nature and type of anger experienced. Low to moderate levels of in-group identification may indicate less fit with the in-group and therefore less importance to oneself perception and identity. When a social identity is less important, threats to that identity will not be seen as serious and self-righteous anger will not be as strong. Righteous or justice-oriented anger, however, may still result.

H1a: Twitter feeds that contain explicit references to a social identity group will result in more self-righteous anger than Twitter feeds that do not contain social identity group references.

H1b: Twitter feeds that contain explicit references to a social identity group will result in less righteous anger than Twitter feeds that do not contain social identity group references.

H1c: Stronger identification with a social identity group will serve as a moderator and will strengthen the relationship of the presence of social identity group references in Twitter and self-righteous anger.

Yet, anger is often induced beyond social identity and under other conditions of social media experience as well. Specifically, the extent to which tweets are polarizing may drive their emotional outcomes as well.

Polarization in Post Content and User Emotion

It is important to consider the nature or content of a user's commentary, as this often guides the direction of responses to a Twitter feed. Specifically, the extent to which the perspectives presented on a Twitter feed are polarizing in discourse and the way the content is presented can exert differential effects on emotions. Kim and Kim (2019) found that polarizing content that was perceived as offensive to one's identity resulted in more negative than positive emotions. Their research supports an interaction effect between the way in which comments are presented and the extent to which commentary is polarizing, such that the more polarizing and incongruent commentary was with one's identity, the more negative emotions were induced relative to positive emotions. In these settings, anger is often a product of this high polarization and incongruence with social identity (Kim & Kim, 2019; Lu & Lee, 2018). However, the means

by which self-righteous relative to righteous anger are affected remains relatively unknown, and social identity is likely to play a role. It is possible that more polarizing Twitter feeds which include explicit references to social identity groups, may result in more self-righteous anger when compared to polarizing feeds that do not contain such information, due to the perceived threat to the in-group and thus one's social identity.

H2a: Twitter feeds that are high in polarization will result in more righteous anger and self-righteous anger than Twitter feeds that are low in polarization.

H2b: Explicit mention of a social identity group will moderate the relationship of Twitter feed polarization to self-righteous anger such that the presence of both social identity information and high polarization will result in more self-righteous anger than Tweets low in polarization with or without social identity group references.

However, polarization and social identity's impact may extend beyond induced anger. The degree to which polarization and social identity may affect one's ability to accurately interpret and recall details of an online debate, in addition to attitudinal stance(s) and strength of attitude(s), may be affected as well.

Social Identity, Polarization and Cognitive Biases

Protecting oneself against information that threatens central aspects of oneself in online forums (i.e., social identity threat) may result in a user being likely to prefer topics and exchanges that align with their in-group's beliefs, values and perspective (Jost & Krochik, 2014). The caveat to this is that information that contradicts a user's perspectives may be intentionally or unintentionally overlooked or dismissed, particularly if a user has experienced negative affect as a product of engagement with online material (Knobloch-Westerwick, Mothes, & Polavin, 2020). This process of selectively awarding attention and/or value to information that aligns with

one's perspective while disregarding contradicting information is known as confirmation bias (Klayman, 1995), and exists across the topic spectrum in online political information gathering and discussion (e.g., Knobloch-Westerwick, Mothes, Johnson, Westewick, & Donsbach, 2015). As a result, online political exchanges often fall victim to confirmation bias, which may lead to a variety of negative outcomes. Regarding social identity and polarization and their effects on one's ability to interpret, recall, and effectively engage with other users' commentary, a variety of relationships may occur.

For example, when content is polarizing, it can increase the strength to which one identifies with the issue(s) or identity-stance discussed (e.g., Mackie, 1986). Content that is polarizing can have the effect of swaying one's opinion from a centrist view or stance to one that is more strongly defined (Prior, 2013). As a result, a user may become biased towards information that emboldens these stances and, in this sense, users may be drawn mostly to others who essentially support their identification with a particular perspective on an issue. The more polarizing the issue, the more users might fall into extremes of supporting one side over the other—or the in-group relative to the out-group. When this occurs, one may have difficulty broadly interpreting information and seeing the argumentative structure in online debates. Attention is then focused on comments that are in-line with social identity expression (i.e., confirmation bias) and disregarding or delegitimizing users whose perspectives may negate the original poster's content. In online forums with political discussion, confirmation bias can affect attitudes and political identification by strengthening the extent to which one feels trust and a personal connection to their identified group (Knobloch-Westerwick, Johnson, & Westerwick, 2015). Information confirming one's perspectives and attitudes strengthens this relationship, whereas information that is noncongruent weakens this relationship (Knobloch-Westerwick,

Johnson, & Westerwick, 2015). As a result, information that is highly polarizing and in which social identity is made highly relevant may decrease depth of information processing and hinder one's ability to accurately recall information and respond to a user's contradicting perspective(s) effectively. This phenomenon may produce a variety of social media interaction outcomes that are ultimately affected by the strength of the attitude expressed by a user and the amount of information they are able to both recall and convey in sharing their perspective online. When controversial topics are discussed in these online interactions, this means recalling perspectives other than that are consistent with one's own. Additionally, as with emotional reactions to Tweets that contain social identity information, strength of social identity will likely strengthen the relationship between explicit social identity references in social media content and confirmation bias.

H3a: Twitter feeds that are high in polarization will lead to more confirmation bias than feeds that are low in polarization.

H3b: Twitter feeds that contain explicit reference to social identity group will lead to more confirmation bias than feeds that do not contain reference to a social identity group, especially when strength of social identification is stronger.

H3c: Twitter feeds that are high in polarization and contain explicit references to social identity group will lead to greater confirmation bias than when only polarization or social identity information is present.

H4a: Twitter feeds that are high in polarization will lead to stronger attitudinal perspectives expressed in user responses than feeds that are low in polarization.

H4b: Twitter feeds that contain high social identity will lead to stronger attitudinal perspectives expressed in user responses than feeds that are low in social identity, especially when strength of social identification is stronger.

H4c: Twitter feeds that are high in polarization and contain explicit references to social identity group will lead to stronger attitudinal perspectives expressed in user responses than when only polarization or social identity information is present.

In a similar vein, positive and negative emotions can also heighten the extent to which one's depth of information processing is either hindered or improved in these environments through aspects of their cognitive appraisal patterns (Stavraki et al., 2021). Anger associated with threats to one's social identity (self-righteous anger) versus anger that reinforces positive self-views (righteous or justice-oriented anger) can lead to decreased argumentative discrimination quality and interpretation which could result in more confirmation bias (Stavraki et al., 2021). While the research that has examined the effects of emotions on information processing regarding argumentative content has established effects from anger in a broader sense, the extent to which anger that is righteous or self-righteous impacts information processing has not been explicitly tested. On one hand, anger overall may increase depth of information processing, as users may be more motivated to seek information that discredits their perspectives as a means to offset feelings of a lack of control over a situation (Moons & Mackie, 2007). Additionally, anger may also impact the chances that one will consider alternative perspectives while enhancing understanding of their original stance (Young, Tiedens, Jung, & Tsai, 2010). That is, the more anger that one feels and the degree the anger is related to their identity, the more one may attempt to understand broader information that could either discredit or support their position. However, anger that is felt to a stronger and perhaps more personal degree (i.e., self-righteous

anger) may result in greater confirmation bias than anger that is more focused on justice for society in general (i.e., righteous anger). As a result, there may be a potential relationship of anger with the observed relationships between social identity and content polarization on confirmation bias. As a product of emotion, there is often the continuation of interaction with a user through responding to their posts. In developing responses and understanding the nuances of these discussions and the points made, emotions may potentially exert impacts on the relationship between one's ability to process information after encountering an online exchange that contains aspects of social identity and/or polarization.

RQ1: How might righteous and self-righteous anger mediate the relationships between social identity and content polarization and attitudinal strength expressed in Twitter responses?

RQ2: How might righteous and self-righteous anger mediate the relationships between social identity and content polarization and confirmation bias?

Other types of cognitive biases might also be important to consider with respect to processing information in online spaces, especially when aspects of one's social identity and content polarization are present in these discussions (Wang, Sirianni, Tang, Zheng, & Fu, 2020). Understanding the nuances of additional cognitive biases in these spaces and how righteous and self-righteous anger may impact the relationship of polarization and social identity to biases such as selective attention, emotional content bias, and commitment bias are important to consider as well.

Selective Attention Bias

Selective attention may bias a user's ability to interpret and interact with others online by narrowing one's focus to aspects of an argument that are congruent with the user's existing

attitudes, while disregarding contradictory information (Bater & Jordan, 2019). In online spaces, Sülflow, Schäfer, and Winter (2019) examined how selective attention may guide preferences towards politically charged news sources on social media (i.e., Facebook) and found that posts with high credibility and reinforcement of users' attitudes resulted in greater selection and attention given to this information over posts that did not reinforce users' attitudes. This research supports the notion that social media users often gravitate towards information that aligns with their personal views (i.e., strength of social identity) and social identity. As a result, the presence of social identity content may exert stronger effects on selective attention bias than the absence of social identity content, as users may potentially award more attention and recall more details that are congruent with their social identity. In a similar vein, when recalling details of a Twitter feed, the strength of a given user's attitudinal perspective and their pre-existing, associated attitudinal stance (on masking) may strengthen this relationship with selective attention bias as well. That is, stronger attitudinal perspectives may moderate the relationship between social identity and selective attention bias, such that the presence of social identity characteristics and stronger attitudinal perspectives will lead to greater selective attention bias than the absence of social identity characteristics or weaker attitudinal perspectives.

H5a: The presence of social identity characteristics will lead to more selective attention bias than the absence of social identity characteristics.

H5b: Strength of social identification with a group will moderate the relationship between explicit references to a social identity group and selective attention bias, such that the presence of social identity information and strength of social identification with a group will lead to greater selective attention bias than the absence of social identity characteristics or weaker strength of social identification with a group.

In addition to social identity and strength of attitudinal perspective exerting effects on selective attention bias, users may also be influenced by the emotional tone and subsequent discrete emotions induced. For example, emotions such as anger that are often present in these discussions may lead to selective attention bias, as Finucane (2011) proposed in their study that explored induced fear and anger's relationships with selective attention. In their research, the author found that anger, as a high arousal emotion, led to more selective attention bias than the control or neutral emotion. However, while further research has examined anger beyond induced emotional states, specifically trait anger (Ford and colleagues, 2012), and found similar trends of biased information processing, research has yet to examine more specific types of anger (i.e., righteous vs. self-righteous). It is possible that righteous anger, which is more focused on injustices that are societally based, may share a weaker relationship with selective attention bias than self-righteous anger, which is more focused on protecting oneself and their inner social circle. Self-righteous anger, which may be more associated with one's social identity and attitudes could lead to a stronger relationship with selective attention bias as one attends to information that is aligned with these facets. Yet, the extent to which righteous and self-righteous anger may act as mediators in the relationships between social identity, strength of attitudinal perspective, and selective attention bias remains more ambiguous and warrants exploration.

RQ3: How might righteous and self-righteous anger mediate the relationships between social identity and strength of social identification with a group on selective attention bias?

Beyond selective attention bias, the effects of emotions such as anger and their emotional tone in online debates can affect information processing as well, particularly when content is polarizing. When users interpret twitter feeds that contain information that may be affective in

nature, this can potentially result in challenges to understanding the many and sometimes complex elements of the debates they engage with online (Bailey, 2021). Subsequently, bias as a product of emotional content may affect engagement with other users in online debate, too.

Emotional Content Bias

Emotional content in online discussions often leads to certain aspects of information becoming more viral or gaining more salience in attention than information that is less emotional in content, particularly when that information evokes high arousal (e.g., anger) emotion in users as a product of interacting with posts (Berger & Milkman, 2012). Indeed, Vochocová, Numerato, and Sedláčková (2022) found that comments in the COVID-19 vaccination online debate that were more polarizing in nature (i.e., extreme views expressed by users) may lead to both stronger dichotomies of perspectives expressed (e.g., pro- vs. anti-vaccine) and emotions experienced by users. The COVID-19 mask debate may follow a similar trend, in which strength of attitudinal perspectives on masking may moderate the relationship between social media post content that is polarizing and its effects on emotional content biasing. It is possible that when users engage with twitter feeds that express more extreme views, this may result in more emotional content bias when recalling aspects of the argument, and the greater the user's attitudinal perspectives central to the discussion (i.e., strength of attitudinal perspective towards masking), the stronger this relationship may manifest in online debates. However, the answer of how righteous and self-righteous anger may mediate these relationships also remains less clear. If users focus on the emotional content of debates, anger may become more salient as these users experience righteous and self-righteous anger from polarization. It is possible that righteous and self-righteous anger may act as additional effects on emotional content bias by producing more focus on emotions perceived as congruent with how a user feels when they are processing information relative to

the debate they are witnessing. As a result, the nature of how these types of anger may act as mediators needs further examination.

H6a: Twitter feed content that is high in polarization will result in more emotional content bias than twitter feed content that is low in polarization.

H6b: The relationship between Twitter feed content polarization and emotional content bias will be moderated by strength of social identification with a group, such that greater social identification with a group will strengthen the relationship between content polarization and emotional content bias.

RQ4: How might righteous and self-righteous anger mediate the respective relationships between polarization, strength of social identification with a group, and emotional content bias?

Similarly, when experiencing emotions in online debate, commitment bias may arise as users are attempting to either support or reject a given perspective: commitment bias. This bias focuses on past behaviors and perspectives that are used to shape future behaviors and perspectives that are consistent with one another, and shares relationships with social identity and content polarization (e.g., Grigoropoulou, 2020). For example, one may justify past research on the effectiveness of masking to prevent the spread of COVID-19 as reason to continue masking in the future to prevent the spread of COVID-19, regardless of contradictory information that may cite evidence that masking may not be as preventative as other techniques (e.g., vaccination).

Commitment Bias

Social media information that is consistent with a user's past behavior and perspectives or attitudes may result in bias towards these attributes for guiding future behaviors and

perspectives or attitudes. In online debate, this relationship has only recently been explored, with literature supporting a sort of consistency bias with these outcomes. Hunzaker (2016) explored how conversations central to aspects of social identity (e.g., cultural characteristics) may shape preferences towards future perspectives and found that information processing is often favored and sometimes altered to be consistent with one's pre-existing characteristics. When information is consistent with pre-existing characteristics of one's social identity, this information may be prioritized or shared to a greater degree than contradictory information to support one's perspective (i.e., an echo chamber) (Garimella and colleagues, 2018). Subsequently, social identity may strengthen the extent to which commitment bias occurs, as individuals may seek information that confirms past behaviors and perspectives or attitudes that are consistent with how they identify in order to guide their recommendations for future courses of action. However, the extent to which polarization may act within these settings is less clear. On one hand, polarization in twitter feeds could result in users strengthening their commitment to past behaviors and perspectives or attitudes to bolster support for their present and future recommendations of actions and cognition as well. Yet, polarization could also potentially weaken this relationship, particularly if one perceives their 'side of the debate' as performing unfavorably in the online discussion. As a result, although polarization may share a relationship with commitment bias, the extent to which it exerts influence on commitment bias is more ambiguous. Additionally, how righteous and self-righteous anger may respectively mediate these relationships has yet to be explored.

H7: The presence of explicit references to a social identity group will result in more commitment bias than the absence of social identity information.

RQ5: How might high and low polarization impact commitment bias?

RQ6: How might righteous and self-righteous anger mediate the respective relationships between social identity, polarization, and commitment bias?

Method

Sample

This research utilized a 2 (social identity info present vs. absent) x 2 (high vs. low polarization) factorial design. Initially, 412 participants were collected. After removing participants who did not complete key portions of the survey, moved through the survey too quickly, or missed the attention checks, the final sample size was 357 participants (i.e., average age = 18.71 years old; 82.4% female; pro-mask = 174, anti-mask = 41, neutral/undecided = 142). Through power analysis with the GPower 3.1 software, it was determined that this sample size was sufficient to proceed with data analytics (Faul, Erdfelder, Buchner, & Lang, 2009). This sample was drawn from undergraduate students in the Department of Psychology at the University of Oklahoma through the online recruitment platform, SONA. These undergraduate students were compensated with course credit for their participation in an online, one-hour maximum study. Following a method similar to Knobloch-Westerwick and colleagues (2020), participants were pre-screened for how long they have lived in the United States. In line with these authors' recommendations, participants who have not lived in the United States since birth may not identify with political affiliations as strongly regarding identity as participants who have experienced the United States' political environment since birth. Subsequently, only participants who met this prerequisite were recruited for this study.

Measures & Materials

The measures for this study included: A simulated twitter feed of the 'mask debate' surrounding COVID-19 preventative measures; an adapted version of Pagano & Huo's (2007)

Moral Outrage Scale that consisted of six pro-mask and six anti-mask moral outrage items (i.e., three righteous anger; three self-righteous anger per scale); the Big Five Inventory (John, Donahue, & Kentle, 1991; John Naumann, & Soto, 2008); a measure of verbal reasoning (Ruch & Ruch, 1980); the Moral Foundations Questionnaire or MFQ (Graham, Haidt, & Nosek, 2008); a 10-item measure of proneness to confirmation bias (Rassin, 2008); a developed measure of confirmation bias as an outcome variable (Ardi & Pradiri, 2021; He et al., 2021; Meppelink et al., 2019); a trait anger scale (Spielberger et al., 1983); and demographics, including questions regarding gender, age, nature and strength of political affiliation (i.e., strength of social identity), and a short scale to identify the participants' stance on masking to prevent the spread of COVID-19. These measures were presented via the online survey platform, Qualtrics.

Procedure

Participants were informed that would have the option to participate in an hour-long online study through Qualtrics that looks at the relationship of social media engagement to psychological outcomes. After consent was confirmed, participants began by first completing a battery of measures in the following order: The MFQ, Big Five Inventory, the proneness to confirmation bias measure, the trait anger scale, and the verbal reasoning measure. After completion of these items, participants were then assessed on their stance on masking to prevent the spread of COVID-19, their social proximity to COVID-19 (e.g., if they or someone in their social circle had COVID-19, vs. no exposure), a political affiliation assessment (select the U.S. political party they identify most with: Democrat (n = 122), Republican (n = 122), Independent (n = 90), or Other (n = 23)) and extent to which they identify with that political party (1 = very little, 2 = somewhat, 3 = moderate extent, 4 = high extent). Upon completion of this portion of the survey, Qualtrics then randomly assigned participants to the four study conditions.

Participants were then exposed to a simulated Twitter feed of 11 tweets discussing whether or not people should wear masks to prevent the spread of COVID-19, with four tweets representing the anti-mask perspective, four tweets representing the pro-mask perspective, and three neutral tweets that neither indicated support nor aversion towards masking. Each condition contained (high/low) polarization and (high/low) social identity relevance (explicit mention of Democrats and Republicans). After reading through this feed, participants then advanced to the next page, where they were given the adapted moral outrage scale measure to assess the extent to which they felt pro-mask and anti-mask righteous and self-righteous anger (adapted from Pagano & Huo, 2007). Then, participants were asked to post a reply comment or tweet to the Twitter feed. Upon posting the comment, they were advanced forward in the survey and asked to recall the original twitter discussion in as much detail as they could. Following this, they completed the adapted confirmation bias outcome measure, a brief measure that assessed the strength of and their stance towards masking, demographics, and were then be debriefed on the nature of the study.

Manipulations

Content of the Twitter feeds were manipulated in this study to reflect (high/low) polarization and (high/low) political social identity characteristics. Similar to the methodology of Mery (2020), the tweets used for this endeavor were created by drawing on actual social media feeds and Tweets about COVID-19 masking and were formatted using an online Twitter feed generator that created a realistic look to the Twitter feed by providing the characteristics of the fictional users (e.g., Twitter handle, profile picture). For the purpose of this research, the tweets for this study were generated from a thorough review and compilation of discussion online through a search of hashtags including but not limited to: #nomaskmandate; #nomasks;

#promask; #antimask; #mandatemasks; and #pleasemask. Based on the social media threads found through these discussions, the simulated Twitter feed discussions were generated. For the polarizing conditions, participants who were exposed to high polarization observed a discussion in which both anti-mask and pro-mask extreme perspectives were explicitly and forcefully expressed, whereas for the low polarization condition, anti-mask and pro-mask perspectives were expressed moderately in a more subdued way. For the social identity conditions, those in the high social identity conditions were exposed to explicit reference to political parties, whereas those in the low social identity condition were exposed to no explicit references to political parties' labels. (Refer to Appendix A through Appendix D for simulated Twitter feeds).

Dependent Variables

The dependent variables assessed in this study included righteous anger; self-righteous anger; stance on and strength of attitude towards masking; and the cognitive biases, including confirmation bias, selective attention bias, emotional content bias, and commitment bias.

Righteous and Self-Righteous Anger

Righteous anger is defined as the perception that there should be justice for victims of a wrongdoing; whereas self-righteous anger is defined as the perception that there should be justice for one's in-group victims of a wrongdoing. Righteous and self-righteous anger were first assessed using a methodology similar to Rothschild and Keefer (2018), through the use of an adapted version of Pagano and Huo's (2007) Moral Outrage Scale.

The adapted Moral Outrage Scale was given to participants after they were exposed to the simulated Twitter feed. This scale consisted of 12-items on a 5-point Likert scale (from *strongly disagree* to *strongly agree*) and included items that measure both types of anger. Half of the items assessed righteous anger, with six statements (3 reflecting the pro-masking stance, and 3

reflecting anti-masking stance), whereas the other half of the items assessed self-righteous anger (3 reflecting the pro-masking stance, and 3 reflecting anti-masking stance). For example, pro-mask righteous anger was captured by the item, "*I am angry about how new surges of COVID-19 could affect society because people will not wear masks.*"; whereas pro-mask self-righteous anger was captured by the item, "*I am angry about how new COVID-19 surges will affect me because others will not wear masks*." Anti-mask righteous anger was captured by the item, "*I am angry about how mask mandates restrict people's freedom*"; whereas anti-mask self-righteous anger was captured by the item, "*I am angry because mask mandates restrict my freedom*." The purpose of these scales was to evaluate the extent to which participants felt righteous and self-righteous anger and was divided into two portions to assess the contrasting perspectives of pro-mask and anti-mask righteous and self-righteous anger, respectively. While the original scale typically assesses one perspective (i.e., 'pro-' or 'anti-') of a debate, this research assessed two different perspectives. Subsequently, a scale with additional items to capture these contrasting relationships was used in the survey (Refer to Appendix E for the adapted Moral Outrage Scale). For scoring of this measure, four sub-scales were produced, including: Pro-mask righteous anger ($\alpha = .90$), pro-mask self-righteous anger ($\alpha = .90$), anti-mask righteous anger ($\alpha = .81$), and anti-mask self-righteous anger ($\alpha = .87$).

Stance and Strength of Attitude Towards Masking

The attitudinal stance towards masking (r*wg = .92) and the strength of this attitude (r*wg = .70) were evaluated as well. The attitudinal stance towards masking is defined as the perspective in the debate that participants express in their tweets. For this portion of the study, three trained raters evaluated the extent to which participants expressed the presence of a perspective on a 3-point Likert scale in their response tweet (i.e., 1 = *Anti-Mask*, 2 = *Neutral,* 3 =

Pro-Mask). In addition to this, participants' response tweets were evaluated on the strength of their stance expressed. Strength of stance is defined as the extent to which the stance expressed in a tweet is salient and was rated on a 5-point Likert scale (1 = Very Weak, 2 = Weak, 3 = Moderate, 4 = Strong, 5 = Very Strong).

Cognitive Biases

Confirmation Bias

Confirmation bias is defined as the tendency to seek out, interpret, attend to, and recall information in a way that confirms or supports one's existing beliefs or values (Nickerson, 1998). A confirmation bias scale ($\alpha = .73$) was developed through: 1) combining methodology from Meppelink and colleagues (2019) and Ardi and Pradiri (2021), respectively, and 2) by utilizing examples of content from He et al. (2021) investigation into the masking conversation of Twitter as guides for development of items for the present research endeavor. The confirmation bias scale assessed the likelihood that participants would share or retweet information relative to the COVID-19 masking debate. This scale included 5 items with 'tweets' about anti-mask information and 5 items with 'tweets' about pro-mask information that were rated on a 5-point Likert scale (i.e., 1 = Very Unlikely to Share, 2 = Unlikely to Share, 3 = Neutral, 4 = Likely to Share, 5 = Very Likely to Share), with examples including, "*The science is inconclusive if wearing a mask prevents the spread of COVID-19*" for anti-mask, and "*Wearing a mask prevents the spread of COVID-19, similar to covering your nose when sneezing to prevent the spread of infections to others*" for pro-mask (Refer to Appendix F for confirmation bias scale).

Selective Attention Bias

Selective attention bias was defined as the tendency to focus on specific argumentative points and cues that are congruent with aspects of one's identity that may bias accuracy of

memory recall (e.g., Bater & Jordan, 2019). The three trained raters evaluated the argument summaries on two elements of selective attention bias, including accuracy of recalled information and recall of unique details. Accuracy of recalled information referred to a count of how often in the argument summaries participants accurately recalled information that pertained to either the pro-mask (r*wg = .70) or anti-mask (r*wg = .73) elements of the Twitter feed they were exposed to; whereas recall of unique pro- and anti-mask details refers to a count of information that was not present in the twitter feed that was presented as apart of the Twitter feed discussion, whether pro-mask (r*wg = .89) or anti-mask (r*wg = .94). Scores for pro-mask and anti-mask selective attention bias were respectively calculated by adding recall of unique details to accuracy of information central to a given perspective recalled by participants, and subtracting the Anti-Mask Selective Attention Score from the Pro-Mask Selective Attention Score, and then adding a constant (i.e., '+ 3', as '-3' was the lowest integer value) to remove negative integer values. Subsequently, the final score was from 0 – 5.67, with lower scores indicating greater attention awarded to anti-mask argument details, and higher scores indicating greater attention awarded to pro-mask argument details.

Emotional Content Bias

Emotional content bias (r*wg = .72) refers to the tendency to focus on emotions (e.g., anger, hostility) and emotional tone expressed (e.g., negative tones) in an argument and was evaluated on a 5-point Likert scale by the three trained raters (i.e., 1 = Very Low or Not at All, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High). Examples of this included when participants focused their argument summary primarily on emotions perceived, such as "anger", or emotional tones, such as "very negative", in the tweets rather than more explicit details of perspectives or content expressed in the simulated Twitter feeds.

Commitment Bias

Finally, commitment bias (r*wg = .70) refers to the tendency to focus on past behavior and decisions to drive future courses of action and perspectives, argument and was evaluated on a 5-point Likert scale by the three trained raters (i.e., 1 = Very Low or Not at All, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High). Examples of this included when participants focused their argument summary primarily on recalling past research or outside of examples of previous behaviors, such as being vaccinated or taking preventative measures for other types of illnesses such as the flu, to justify future directions for moving forward, such as masking or getting vaccinated to prevent the spread of COVID-19, rather than more explicit details of perspectives or content expressed in the simulated Twitter feeds.

Covariates

Covariates for this study consisted of the Big Five Personality Traits, Trait Anger, progressiveness ideology (i.e., the MFQ), proximity to COVID-19, proneness to confirmation bias, and verbal reasoning. First, it was important to assess the Big Five Personality Traits, including Openness to Experience (α = .81), Conscientiousness (α = .75), Extraversion (α = .88), Agreeableness (α = .79), and Neuroticism (α = .80), as traits such as Conscientiousness and Agreeableness may affect the extent to which participants are able to identify aspects of the Twitter conversation(s) and their ability to interact, particularly if the participants experience induced with anger (e.g., Jensen-Campbell et al., 2007; Pfeiler et al., 2018; Sindermann et al., 2018). Next, Trait Anger (α = .87) was important to consider as well, as Trait Anger may potentially impact the extent to which users engage with one another in online spaces (Wang et al., 2017). Political ideology (i.e., progressiveness, assessed through the Moral Foundations Questionnaire) shares a relationship with political affiliation and may share additional insights

with social identification (Greene, 2004). For analyses, only the subscales of Harm (α = .58), Fairness (α = .62), and Ingroup Loyalty (α = .71) were primarily used to focus on capturing this construct, as Purity and Authority were not the focus of the COVID-19 masking debate discussions presented in the simulated Twitter feeds. However, due to low reliabilities for the Harm and Fairness subscales, only Ingroup Loyalty was used as a covariate measure for this construct in the tested models. Next, proximity to COVID-19 was also explored, as participants who have experienced COVID-19 either personally (i.e., exposure to self or others in social circle) or not all (i.e., no exposure to self or others in social circle) could potentially produce stronger responses to the masking debate discussion. Proneness to confirmation bias (α = .73) was also explored, as this may predict the likelihood that one will be susceptible to confirmation bias (Rassin, 2008). Pre-existing attitudes towards masking could also potentially exert effects on these outcomes as well, such that they may produce stronger effects, particularly for those who have pre-existing beliefs about COVID-19's legitimacy and need for preventative measures (e.g., Georgiou, Delfabbro, & Balzan, 2020). Finally, verbal reasoning was assessed, as participants who rate lower on verbal reasoning may hold differing reactions to COVID-19 than participants who rate higher on verbal reasoning (Batty, Deary, & Gale, 2021).

Data Analysis Approach

The purpose of this research was to assess the relationships between righteous and self-righteous anger, social identity, polarization, and cognitive biases (See Table 1 for means, standard deviations, and correlations among all dependent variables and covariates). In line with recommendations by Cho and Abe (2013), one-tailed tests were utilized for directional predictions and two-tailed tests were utilized for nondirectional predictions and research questions. For hypotheses and research questions specific to moderation analysis with categorical

moderators and direct effects, ANCOVA was utilized to test these relationships, whereas for mediation analyses and moderation analyses with continuous moderators, Hayes's (2022) PROCESS macro was utilized to test these relationships in the statistical software package SPSS v28. Additionally, in line with recommendations by Bernerth and Aguinis (2016), covariates included in each statistical model had theoretical justification for potential relationships, empirical support through prior research, and reliable measurement in this study. Initially, all covariates were included in the models for the assessed relationships with the dependent variables, respectively, and only covariates that did not demonstrate statistically supported relationships were removed from the final models used for analyses of hypotheses and research questions.

Results

Hypotheses Tests

The following tests evaluate the hypotheses with independent samples t-tests (See Tables 2 and 3), ANCOVAs for categorical moderators (See Tables 4, 5, and 6), and the PROCESS macro for continuous moderators (See Table 7). H1a proposed that Twitter feeds that contain references to social identity will result in more self-righteous anger than Twitter feeds that do not contain references to social identity. In order to assess this relationship, an independent samples t-test was used to evaluate pro-mask self-righteous anger and anti-mask self-righteous anger, respectively. Findings indicate that while there is no support for anti-mask self-righteous anger, there is support for pro-mask self-righteous anger, such that the presence of social identity characteristics in Twitter feeds (M = 3.02, SD = 1.27) results in greater self-righteous anger than the absence of social identity characteristics in Twitter feeds (M = 2.75, SD = 1.20), $t(355) = -2.03$, $p < .05$ (see Table 2) Subsequently, there is partial support for H1a. Next, H1b proposed

that Twitter feeds that contain references to social identity will result in less righteous anger than Twitter feeds that do not contain references to social identity. In order to assess this relationship, an independent samples t-test was used to evaluate pro-mask righteous anger and anti-mask righteous anger, respectively. Findings indicate that there is no support for this relationship with either anti-mask righteous anger or pro-mask righteous anger (see Table 2). Subsequently, there is no support for H1b. Finally, H1c proposed that stronger identification with political social identity will serve as a moderator of the relationship between the presence (or absence) of social identity characteristics in Twitter feeds and self-righteous anger. Utilizing PROCESS, while covariate effects were present in both models through moderation analysis that assessed anti-mask and pro-mask self-righteous anger, respectively, including neuroticism ($CI_{.95}$ = .08, .37), openness to experience ($CI_{.95}$ = .04, .36), Pre-Stance Attitude on Masking ($CI_{.95}$ = .80, 1.12), and proximity to COVID-19 ($CI_{.95}$ = -.20, -.00), for pro-mask self-righteous anger, and only ingroup loyalty ($CI_{.95}$ = .21, .47) and Pre-Stance Attitude on Masking ($CI_{.95}$ = -1.08, -.77) for anti-mask self-righteous anger, neither model had statistical support for the proposed moderation effect (see Table 6, models 1 and 2). Subsequently, there is no support for H1c.

H2a proposed that Twitter feeds that contain high polarization will result in more righteous and self-righteous anger than Twitter feeds that contain low polarization. In order to assess this relationship, an independent samples t-test was used to evaluate pro-mask and anti-mask righteous anger and self-righteous anger, respectively. Findings indicate that there is no support for this relationship with either anti-mask or pro-mask righteous anger or self-righteous anger (see Table 3). Subsequently, there is no support for H2a. Next, H2b proposed that explicit references to social identity will moderate the relationship between content polarization and self-righteous anger, such that the presence of social identity characteristics and high polarization will

result in greater self-righteous anger than the presence or absence of social identity information with low polarization. Utilizing ANCOVA, significant covariate effects were present in both models that assessed anti-mask and pro-mask self-righteous anger, respectively. Specifically, significant covariates for anti-mask self-righteous anger included ingroup loyalty, $F(1, 351) = 27.13$, $p < .001$, $\eta 2 = .072$ and pre-stance attitude on masking, $F(1, 351) = 133.71$, $p < .001$, $\eta 2 = .276$; whereas significant covariates for pro-mask self-righteous anger included neuroticism, $F(1, 349) = 9.58$, $p < .05$, $\eta 2 = .026$, openness to experience, $F(1, 349) = 5.98$, $p < .05$, $\eta 2 = .017$, and proximity to COVID-19, $F(1, 349) = 4.17$, $p < .05$, $\eta 2 = .012$ (see Table 4). However, neither model demonstrated statistical support for the proposed moderation effect. Subsequently, there is no support for H2b.

H3a proposed that Twitter feeds that contain high polarization will result in more confirmation bias than Twitter feeds that contain low polarization. In order to assess these relationships, an independent samples t-test was utilized. While there is statistical support for a relationship between confirmation bias and polarization, the findings are opposite of what was predicted. Specifically, Twitter feeds that contain low polarization ($M = 2.48$, $SD = .63$) resulted in more confirmation bias than Twitter feeds that contain high polarization ($M = 2.31$, $SD = .76$), $t(355) = 2.30$, $p < .05$.(see Table 3). Subsequently, there is no support for H3a. H3b proposed that Twitter feeds that contain social identity will lead to more confirmation bias than Twitter feeds that do not contain social identity, especially when strength of social identification with a group is higher. Utilizing PROCESS, a significant covariate effect of verbal reasoning on confirmation bias ($CI_{.95} = -.03, -.01$) was found; however, the findings do not support the proposed moderated relationship between social identity, strength of social identification with a group, and confirmation bias (see Table 7, model 4). Subsequently, there is no support for H3b.

H3c proposed that Twitter feeds that contain high polarization and the presence of social identity characteristics will result in more confirmation bias than Twitter feeds that contain low polarization and the absence of social identity characteristics. While the findings do indicate through ANCOVA a main effect of polarization on confirmation bias, such that Twitter feeds that contain low polarization (M = 2.48, SD = .63) result in more confirmation bias than Twitter feeds that contain high polarization (M = 2.31, SD = .76), $F(1, 351) = 6.16$, $p < .05$, $\eta 2$, = .017, in addition to a covariate effect of verbal reasoning on confirmation bias, $F(1, 351) = 12.36$, $p <$ .001, $\eta 2$, = .040, there is no support for the proposed interaction relationship (see Table 6). Subsequently, there is no support for H3c.

H4a proposed that Twitter feeds that contain high polarization will result in stronger attitudinal perspectives expressed in user responses than Twitter feeds that contain low polarization. Utilizing an independent sample t-test, the findings support this prediction, such that high polarization (M = 2.63, SD = .87) results in stronger attitudinal perspectives expressed in user responses than low polarization (M = 2.46, SD = .88), $t(355) = -1.87$, $p < .05$ (see Table 3) Subsequently, there is support for H4a. H4b proposed a moderated relationship between social identity and strength of attitudinal perspectives expressed in user responses by importance of political social identity, such that Twitter feeds that contain social identity characteristics will result in stronger attitudinal perspectives expressed in user responses than Twitter feeds that contain no social identity characteristics, especially when strength of political identification with a social group is strong. Utilizing the PROCESS macro, while there is a significant covariate effect of pre-stance attitude on masking (CI.95 = .11, .38) on strength of attitudinal perspectives expressed in user responses, there is no support for the proposed moderation effect (see Table 7, model 3). Subsequently, there is no support for H4b. H4c proposed that high polarization and

presence of social identity characteristics will result in stronger attitudinal perspectives expressed in user responses than Twitter feeds that contain low polarization or no social identity characteristics. Utilizing ANCOVA, while there is a significant covariate effect of pre-stance attitude on masking, $F(1, 352) = 12.80$, $p < .001$, $\eta 2 = .035$, on strength of attitudinal perspectives expressed in user responses, there is no support for the proposed relationship (see Table 5). Subsequently, there is no support for H4c.

H5a proposed that the presence of social identity characteristics in Twitter feeds will lead to more selective attention bias than the absence of social identity characteristics in Twitter feeds. Findings indicate no support for H5a (see Table 2). H5b proposed that strength of political identification with a social group will moderate the relationship between social identity and selective attention bias, such that stronger political identification and the presence of explicit references to social identity social identity will lead to more selective attention bias than weaker political attitudes and the absence of explicit references to social identity. Utilizing the PROCESS macro, while there are significant covariate effects of pre-stance attitude on masking ($CI_{.95} = .38, .50$) and proximity to COVID-19 ($CI_{.95} = -.16, -.02$) on selective attention bias, there is no statistical support for the proposed moderation relationship (see Table 7, model 5). Subsequently, there is no support for H5b.

H7a proposed that Twitter feeds that contain high polarization will result in more emotional content bias than Twitter feeds that contain low polarization. Findings indicate support for H7a, such that high polarization ($M = 1.81$, $SD = .58$) results in more emotional content bias than Twitter feeds that are low polarization ($M = 1.59$, $SD = .44$), $t(355) = -4.07$, $p < .001$ (see Table 3). H7b proposes that the relationship between polarization and emotional content bias will be moderated by strength of political social identity, such that stronger political identification

with a social group will strengthen the relationship between content polarization and emotional content bias. Utilizing the PROCESS macro, findings indicate a significant main effect of polarization ($CI_{.95}$ = .06, .70) on emotional content bias, in addition to a significant covariate effect of neuroticism ($CI_{.95}$ = .05, .19) on emotional content bias (see Table 7, model 6). However, the findings do not support this hypothesis.

Finally, H8 proposed that the presence of social identity characteristics will result in more commitment bias than the absence of social identity characteristics. This relationship was assessed utilizing an independent samples t-test. Findings indicate a lack of support for H8 (see Table 2).

Research Questions Tests

The following tests evaluate the research questions with the PROCESS macro to evaluate moderated mediation relationships (See Table 7). RQ1 explored how pro-mask and anti-mask righteous and self-righteous anger, respectively, may mediate the relationship between polarization, social identity, and strength of attitudinal perspectives expressed by users in their response tweets. For pro-mask angers, findings indicate a direct effect of polarization on strength of attitude ($CI_{.95}$ = .01, .37) (see model 1); however, neither pro-mask righteous nor self-righteous anger mediate the relationships between polarization, social identity, and strength of attitudinal perspectives. Additionally, there are also neither mediation relationships shared with anti-mask righteous or self-righteous anger on strength of attitudinal perspectives, nor direct effect relationships shared with other predictors (see model 2). Subsequently, neither righteous nor self-righteous anger mediate the relationship between polarization, social identity, and strength of attitudinal perspectives expressed by users in their response tweets.

RQ2 explored how pro-mask and anti-mask righteous and self-righteous anger, respectively, may mediate the relationship between polarization, social identity, and confirmation bias in argumentative summaries. Content polarization demonstrated a direct effect on confirmation bias ($CI_{.95}$ = -.30, -.02), in addition to verbal reasoning ($CI_{.95}$ = -.03, -.01) also demonstrating a relationship with confirmation bias when both types of pro-mask anger were included as potential mediators (see model 3). However, no mediation effects were found. When anti-mask angers were included in the model as potential mediators, polarization ($CI_{.95}$ = -.32, -.04) and verbal reasoning ($CI_{.95}$ = -.03, -.01) demonstrated direct relationships with confirmation bias, but neither righteous nor self-righteous anger mediated the relationships between polarization, social identity, and confirmation bias (see model 4). Subsequently, neither righteous nor self-righteous anger mediate the relationship between polarization, social identity, and confirmation bias.

RQ3 explored whether pro-mask and anti-mask righteous and self-righteous anger, respectively, mediate the relationships between social identity and strength of political social identity on selective attention bias. The findings indicate agreeableness ($CI_{.95}$ = -.29, -.06) and pre-stance attitude on masking ($CI_{.95}$ = .05, .34) demonstrated direct relationships with selective attention bias for pro-mask anger analyses (see model 5). Similarly, agreeableness ($CI_{.95}$ = -.29, -.05) and pre-stance attitude on masking ($CI_{.95}$ = .16, .44) demonstrated relationships with selective attention bias for anti-mask angers (see model 6). Neither righteous nor self-righteous anger acted as mediators for the relationship between social identity, strength of political social identity, and selective attention bias.

RQ4 explored how pro-mask and anti-mask righteous and self-righteous anger, respectively, may mediate the relationships between polarization and strength of political social

identity on emotional content bias. While polarization ($CI_{.95}$ = .11, .33) and neuroticism ($CI_{.95}$ = .05, .20) exert direct effects on emotional content bias, neither pro-mask righteous nor self-righteous anger mediated these relationships (see model 7). Similarly, when evaluating anti-mask angers as mediators, polarization ($CI_{.95}$ = .12, .33) and neuroticism ($CI_{.95}$ = .04, .19) exerted direct effects on emotional content bias, but neither anti-mask righteous nor self-righteous anger mediated these relationships (see model 8). There were no direct, indirect, or mediated relationships for strength of political social identity and emotional content bias. Subsequently, neither righteous nor self-righteous anger mediate the relationship between polarization, strength of political social identity, and emotional content bias.

RQ5 explored whether high vs. low Twitter feed content polarization exerts effects on commitment bias. Through independent samples t-test, it was determined that there is no support for this relationship (see Table 3). Subsequently, polarization does not influence commitment bias.

Finally, RQ6 explored how pro-mask and anti-mask righteous and self-righteous anger, respectively, may mediate the relationships between polarization and social identity on commitment bias. Findings indicated that no direct, indirect, nor mediated effects were observed (see models 9 and 10). Thus, neither righteous nor self-righteous anger mediate the relationship between polarization, social identity, and commitment bias.

Discussion

While these findings have some exciting implications and future directions for research, it is important to first address the limitations of this study. First, the sample for this research primarily consists of pro-mask and neutral/undecided participants regarding self-report attitudes on masking to prevent the spread of COVID-19, with a much smaller portion of the sample

consisting of participants who self-reported as anti-mask. While masking stance was included as a covariate, this limitation affects the generalizability of the results of this study to draw inferences about anti-masking attitudes and online users. As a result, interpretations and future directions should take this limitation into consideration when determining recommendations about how to approach conversations of preventative measures to COVID-19 in online spaces. However, it is important to also note that because this sample consists of mostly pro-mask and neutral/undecided participants, findings may be more generalizable to these audiences and may share insight into how online users who identify within these groups may fall victim to emotional and cognitive biases effects when engaging in debate with users of contrasting perspectives. Utilizing the findings of this study may provide additional guidance to these individuals regarding how to interact with others online and both increase cognizance and decrease susceptibility to biasing of information processing when engaging in online debate around contentious topics, such as masking to prevent the spread of COVID-19. Further investigation could potentially alleviate this concern by recruiting a broader sample of participants to ensure representativeness of the sample with regard to pro- and anti-mask perspectives.

Second, while the simulated Twitter feeds were reflective of observed trends of actual conversations on Twitter surrounding this topic of debate, it is important to highlight that political identity is a very complex and multifaceted aspect of social identity and by focusing on primarily Republicans and Democrats in the simulated discussions, there is the potential of both excluding participants of other political party affiliation identities (e.g., Independent, Tea Party) and ideologies unique to political social identity (e.g., Liberal, Conservative, Moderate). In fact, 31.6% of the sample included in this study was comprised of participants who either identified as an Independent or Other political affiliation (i.e., their political social identity). Thus, effects of

the manipulation of social identity in the Twitter Feed and strength of social identity may be underestimated given the inclusion of these individuals in the sample. That is, it is possible as well that participants who identified as Independent/Other may not have experienced induction of the political social identity manipulation to the same extent as participants who identified as Democrat or Republican. Finally, regarding political social identity, by including progressiveness of ingroup loyalty as a covariate in analyses, it is possible that this may impact the variance associated with nuances of political ideology and its relationships with political social identity, which may explain the extent to which ingroup loyalty accounted for much of the variance in its tested models. Accordingly, future research could control for this limitation by representing more diversified political party affiliation and political ideologies, respectively, in the simulated Twitter feeds, or by pre-screening and including only participants with specific political identities.

Regarding the experimental design, there are also limitations to the number of hypotheses and tests and other characteristics of the sample itself. First, there were many predictions made based off theory and extrapolation of potential relationships. While this provides support for hypothetical directions of these predictions, the considerable number of hypotheses and research questions may have led to the smaller effect sizes, as many tests were conducted and, as a result, the extent to which statistical relationships and inferences could be drawn from the sample were significantly diminished.

Additionally, the nature of the sample is of concern as well. First, within the collegiate age group used in this study, Twitter has over years declined in popularity, with platforms such as Instagram and Snapchat rising to lead social media expressions and engagement among the present sample's collegiate age demographics (e.g., Alsalem, 2019). It is possible that Twitter

may not be as relevant to this sample and, as a result, engagement may not have been as effective as use of simulated Instagram feeds—which allows a different type of communication expression (e.g., more character-content). Additionally, many of these students may be habituated to these types of discussions on Twitter (which may have been used more frequent prior to Instagram/when the participants were younger in age) as many political icons often express their perspectives through Twitter posts. Subsequently, these young adults may not be as susceptible to manipulations of political social identity and content polarization on Twitter due to familiarity with this type of content on Twitter over time. Future research could perhaps use an older sample of participants, which may potentially yield more insight into engagement with Twitter and these types of conversations. Finally, the topic of COVID-19 itself may be exhausted within these settings, as the pandemic, a (traumatic) historical event for the last two years, may have lost much of its recency and may result in different or weaker reactions than other contentious topics often expressed on social media that are more common, enduring concerns over expansive time.

Lastly, this experiment focused on anger as an emotional outcome alone and sought to extend the literature on righteous and self-righteous anger as a result. However, these debates often produce additional emotions such as fear (e.g., Li & Xiao, 2020), which may act as an additional impact on the nature of how online users engage with one another. It is possible that discussing COVID-19, a topic that is often surrounded by different drivers of fear (Schimmenti, Billieux, Starcevic, 2020), may also manifest as righteous or self-righteous—a relationship that has yet to be examined in the literature. Subsequently, future research could perhaps examine fear within these constructs and 1) explore if there is support for righteous and self-righteous fear, and 2) contrast such findings to righteous and self-righteous anger. Despite these limitations, this study produced some intriguing findings.

This research observed several relationships between polarization, social identity, righteous and self-righteous anger, and contrasting cognitive biases that unmask how online users engage with each other in online debate. Regarding polarization, it is interesting that the different forms of righteous and self-righteous anger explored were not affected by polarization in the Twitter feed. This finding contradicts previous research that proposes more specific types of anger (e.g., incidental anger; Huber, Van Boven, Park, & Pizzi, 2015) share causal relationships with polarization and politicization. This finding also may have occurred for contrasting reasons, as it is possible that while participants experienced anger, this anger may not have been unique to content polarization, as the findings do support that the presence (or absence) of social identity characteristics increased pro-mask self-righteous anger. This finding suggests certain types of anger may be products of contentious discussions that reference political social identities. While social identity content did not affect pro-masking righteous anger and anti-masking righteous and self-righteous anger, it is possible that social identity content in these kinds of debates is threatening to the self-concept thereby triggering self-righteous anger as a defense. The much larger numbers of pro-mask and neutral/undecided (vs. anti-mask) participants in this sample could explain why this emerged only for pro-mask self-righteous anger. This warrants additional investigation. As previously suggested, balancing recruitment of participants would be helpful to overcome this limitation.

Next, content polarization exerted effects on strength of attitudinal perspectives expressed in users' responses and on cognitive biases, including confirmation bias and emotional content bias. These relationships support literature that proposes that polarization can lead to both stronger attitudes and focus on emotions in online debates (e.g., Vochocová, Numerato, & Sedláčková, 2022), which also provides additional support for the observed relationships

between content polarization as direct effect on emotional content bias. Similarly, it was found that confirmation bias is higher in low polarization relative to high polarization Twitter feeds. It may be the case that when one is exposed to high polarization, which often results in more distinct views expressed in an online debate, that this makes these perspectives more salient to the user—thus, mitigating the likelihood of only interpreting information that reinforces their perspective in the online debate.

Lastly, of additional interest is the consistency of covariates across statistical models tested. Ingroup loyalty, verbal reasoning, proximity to COVID-19, pre-stance attitude on masking, and different aspects of the Big Five Personality Traits demonstrated relationships with the outcome variables, regardless of pro-making and anti-masking perspectives and outcomes, which supports existing literature that proposes these relationships should be both considered and accounted for in future experimental design.

Future research could also apply the findings of this study to multiple avenues in the Emotions and online debate literatures. It would be interesting to explore if righteous and self-righteous emotion extends beyond anger to additional negative discrete emotions, and to explore if polarization and social identity, respectively, exert influences on the extent to which they are felt in online spaces. In a similar vein, the findings with emotional content bias are fascinating and warrant more in-depth analysis, as it would be interesting to compare positive emotional tone to negative emotional tone online debates and to see if online users are more susceptible to one area of bias more so than the other. Furthermore, assessing which discrete emotions are singled-out in summarizing these arguments and comparing their frequency may provide insight regarding which types of discrete emotions users may be more susceptible to falling victim to when information processing is biased. Finally, it would be interesting to replicate this study

with online discussions of the COVID-19 vaccine debate, which is both highly contentious and rooted in multiple facets of social identity (e.g., Latkin, Dayton, Yi, Colon, Kong, 2021). It is possible that while masking to prevent the spread of COVID-19 is both important and polarizing, that the COVID-19 vaccination discussion, which is debatably more contentious and controversial, may lead to dissimilar (and potentially stronger) outcomes.

The purpose of this research was to unmask the effects of polarization, social identity, and different types of anger on information processing and attitudes towards masking to prevent the spread of COVID-19 and, while this research sheds light on these relationships, there is still much to explore. The world continues to change as a product of COVID-19, but as we move forward it is essential to remember that as one approaches the Shakespearian-spin on the debate of… "To mask… or not to mask?!" to consider all perspectives, remain aware of biases, and to follow best health practices.

References

Al-Ramahi, M., Elnoshokaty, A., El-Gayar, O., Nasralah, T., & Wahbeh, A. (2021). Public Discourse Against Masks in the COVID-19 Era: Infodemiology Study of Twitter Data. *JMIR Public Health and Surveillance*, *7*(4), e26780.

Alsalem, F. (2019). Why do they post? Motivations and uses of snapchat, instagram and twitter among kuwait college students. *Media Watch*, *10*(3), 550-567.

Ardi, R., & Pradiri, A. P. (2021). Determinant factors of partisans' confirmation bias in social media. *Humanitas Indonesian Psychological Journal*, *18*(1), 1-13.

Bailey, E. A. (2021). Crosstalk: Emotion and Misunderstanding in Online Debate. In *Political Participation on Social Media* (pp. 103-123). Palgrave Macmillan, Cham.

Bater, L. R., & Jordan, S. S. (2020). Selective attention. *Encyclopedia of personality and individual differences*, 4624-4628.

Batty, G. D., Deary, I. J., & Gale, C. R. (2021). Pre-pandemic cognitive function and COVID-19 mortality: prospective cohort study. *European journal of epidemiology*, 1-6.

Bennett, W. L. (2012). The personalization of politics: Political identity, social media, and changing patterns of participation. *The annals of the American academy of political and social science*, *644*(1), 20-39.

Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of marketing research*, *49*(2), 192-205.

Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, *69*(1), 229-283.

Bhasin, T., Butcher, C., Gordon, E., Hallward, M., & LeFebvre, R. (2020). Does Karen wear a mask? The gendering of COVID-19 masking rhetoric. *International Journal of Sociology and Social Policy* (40), 929-937.

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. *Social identity: Context, commitment, content*, 35-58.

Chen, G. M., & Ng, Y. M. M. (2017). Nasty online comments anger you more than me, but nice ones make me as happy as you. *Computers in Human Behavior*, *71*, 181-188.

Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Hołyst, J. A. (2011). Collective emotions online and their influence on community life. *PloS one*, *6*(7), e22207.

Cho, H. C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate?. *Journal of Business Research*, *66*(9), 1261-1266.

Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011, July). Political polarization on twitter. In *Fifth international AAAI conference on weblogs and social media*.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Finucane, A. M. (2011). The effect of fear and anger on selective attention. *Emotion*, *11*(4), 970.

Ford, B. Q., Tamir, M., Gagnon, S. A., Taylor, H. A., & Brunyé, T. T. (2012). The angry spotlight: Trait anger and selective visual attention to rewards. *European Journal of Personality*, *26*(2), 90-98.

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2017, February). Reducing controversy by connecting opposing views. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (pp. 81-90).

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018, April). Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference* (pp. 913-922).

Georgiou, N., Delfabbro, P., & Balzan, R. (2020). COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and individual differences*, *166*, 110201.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Spassena, K., & Ditto, P. H. (2008). Moral foundations questionnaire. *Journal of Personality and Social Psychology*.

Green, J. D., Sedikides, C., Van Tongeren, D. R., Behler, A. M., & Barber, J. M. (2019). Self-enhancement, righteous anger, and moral grandiosity. *Self and Identity*, *18*(2), 201-216.

Greene, S. (2004). Social identity theory and party identification. *Social Science Quarterly*, *85*(1), 136-153.

Gregory, K. (2018). Anger in academic Twitter: Sharing, caring, and getting mad online. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *16*(1), 176-193.

Grigoropoulou, N. (2020). Discussing God: The Effect of (Ir) Religious Identities on Topic-Sentiment Polarization in Online Debates. *Review of Religious Research*, *62*(4), 533-561.

Gruzd, A. (2012). Examining Polarization in Political Social Media: A Case of Twitter and the 2011 Canadian Federal Election. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*.

Guberman, J., Schmitz, C., & Hemphill, L. (2016, February). Quantifying toxicity and verbal violence on Twitter. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (pp. 277-280).

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

He, L., He, C., Reynolds, T. L., Bai, Q., Huang, Y., Li, C., ... & Chen, Y. (2021). Why do people oppose mask wearing? A comprehensive analysis of US tweets during the COVID-19 pandemic. *Journal of the American Medical Informatics Association*, *28*(7), 1564-1573.

Hogg, M. A. (2004). Social Categorization, Depersonalization, and Group Behavior. In M. B. Brewer & M. Hewstone (Eds.), *Self and social identity* (pp. 203–231). Blackwell Publishing.

Hong, S., & Kim, S. H. (2016). Political polarization on twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, *33*(4), 777-782.

Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and personality psychology compass*, *2*(1), 204-222.

Huber, M., Van Boven, L., Park, B., & Pizzi, W. T. (2015). Seeing red: Anger increases how much Republican identification predicts partisan attitudes and perceived polarization. *PLoS one*, *10*(9), e0139193.

Hunzaker, M. F. (2016). Cultural sentiments and schema-consistency bias in information transmission. *American Sociological Review*, *81*(6), 1223-1250.

Jensen-Campbell, L. A., Knack, J. M., Waldrip, A. M., & Campbell, S. D. (2007). Do big five personality traits associated with self-control influence the regulation of anger and aggression?. *Journal of research in personality*, *41*(2), 403-424.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *Big Five Inventory (BFI)* [Database record]. APA PsycTests.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues.

Jost, J. T., & Krochik, M. (2014). Ideological differences in epistemic motivation: Implications for attitude structure, depth of information processing, susceptibility to persuasion, and stereotyping. In *Advances in motivation science* (Vol. 1, pp. 181-231). Elsevier.

Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European journal of communication*, *32*(3), 257-273.

Kim, Y., & Kim, Y. (2019). Incivility on Facebook and political polarization: The mediating role of seeking further comments and negative emotion. *Computers in Human Behavior*, *99*, 219-227.

Knobloch-Westerwick, S., Johnson, B. K., & Westerwick, A. (2015). Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, *20*(2), 171-187.

Knobloch-Westerwick, S., Mothes, C., & Polavin, N. (2020). Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research*, *47*(1), 104-124.

Knobloch-Westerwick, S., Mothes, C., Johnson, B. K., Westerwick, A., & Donsbach, W. (2015). Political online information searching in Germany and the United States: Confirmation bias, source credibility, and attitude impacts. *Journal of Communication*, *65*(3), 489-511.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of learning and motivation*, *32*, 385-418.

Latkin, C. A., Dayton, L., Yi, G., Colon, B., & Kong, X. (2021). Mask usage, social distancing, racial, and gender correlates of COVID-19 vaccine intentions among adults in the US. *PloS one*, *16*(2), e0246970.

Mackie, D. M. (1986). Social identification effects in group polarization. *Journal of Personality and Social Psychology*, *50*(4), 720-728.

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of health communication*, *24*(2), 129-140.

Mery, M. (2020). The Impact of Polarization, Empathy and Topic Salience in Twitter Feeds on Ethical Decision-Making. Thesis ShareOk.org

Moons, W. G., & Mackie, D. M. (2007). Thinking straight while seeing red: The influence of anger on information processing. *Personality and Social Psychology Bulletin*, *33*(5), 706-720.

Lau, V. W., Bligh, M. C., & Kohles, J. C. (2020). Leadership as a Reflection of Who We Are: Social Identity, Media Portrayal, and Evaluations of Hillary Clinton in the 2016 US Presidential Election. *Sex Roles*, *82*(7), 422-437.

Li, J., & Xiao, L. (2020). Emotions in online debates: Tales from 4Forums and ConvinceMe. *Proceedings of the Association for Information Science and Technology*, *57*(1), e255.

Lu, Y., & Lee, J. K. (2019). Partisan information sources and affective polarization: Panel analysis of the mediating role of anger and fear. *Journalism & Mass Communication Quarterly*, *96*(3), 767-783.

Nigmatullina, K., & Bodrunova, S. S. (2018, October). Patterns of emotional argumentation in Twitter discussions. In *International Conference on Internet Science* (pp. 72-79). Springer, Cham.

Pagano, S., & Huo, Y. (2007). The Role of Moral Emotions in Predicting Support for Political Actions in Post-War Iraq. *Political Psychology, 28*(2), 227-255.

Pfeiler, T. M., Weber, H., & Kubiak, T. (2018). Experiencing anger in a social interaction: The role of personality. *Personality and Individual Differences, 132,* 45–51

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, *16*, 101-127.

Rassin, E. (2008). Individual differences in the susceptibility to confirmation bias. *Netherlands journal of Psychology*, *64*(2), 87-93.

Roy, S., & Ghosh, P. (2021, May). A Comparative Study on Distancing, Mask and Vaccine Adoption Rates from Global Twitter Trends. In *Healthcare* (Vol. 9, No. 5, p. 488). Multidisciplinary Digital Publishing Institute.

Rothschild, Z. K., & Keefer, L. A. (2018). Righteous or self-righteous anger? Justice sensitivity moderates defensive outrage at a third-party harm-doer. *European Journal of Social Psychology*, *48*(4), 507-522.

Ruch, F. L., & Ruch, W. W. (1980). *Employee aptitude survey*. Los Angeles, CA: Psychological Services.

Sanders AC, White RC, Severson LS, Ma R, McQueen R, Alcântara Paulo HC, Zhang Y, Erickson JS, Bennett KP. Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse. AMIA Jt Summits Transl Sci Proc. 2021 May 17;2021:555-564. PMID: 34457171; PMCID: PMC8378598.

Schimmenti, A., Billieux, J., & Starcevic, V. (2020). The four horsemen of fear: An integrated model of understanding fear experiences during the COVID-19 pandemic. *Clinical Neuropsychiatry*, *17*(2), 41.

Sindermann, C., Luo, R., Zhao, Z., Li, Q., Li, M., Kendrick, K. M., ... & Montag, C. (2018). High ANGER and low agreeableness predict vengefulness in German and Chinese participants. *Personality and Individual Differences*, *121*, 184-192.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social psychology quarterly*, 224-237.

Stavraki, M., Lamprinakos, G., Briñol, P., Petty, R. E., Karantinou, K., & Díaz, D. (2021). The influence of emotions on information processing and persuasion: A differential appraisals perspective. *Journal of experimental social psychology*, *93*, 104085.

Sülflow, M., Schäfer, S., & Winter, S. (2019). Selective attention in the news feed: An eye-tracking study on the perception and selection of political news posts on Facebook. *new media & society*, *21*(1), 168-190.

Tajfel, H., & Turner, J. C. (1979). An intergrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tamburrini, N., Cinnirella, M., Jansen, V. A., & Bryden, J. (2015). Twitter users change word usage according to conversation-partner social identity. *Social Networks*, *40*, 84-89.

Trepte, S., & Loy, L. S. (2017). Social identity theory and self-categorization theory. *The international encyclopedia of media effects*, 1-13.

Tripp, T. M., & Bies, R. J. (2010). "Righteous" anger and revenge in the workplace: The fantasies, the feuds, the forgiveness. In *International handbook of anger* (pp. 413-431). Springer, New York, NY.

Vochocová, L., Numerato, D., & Sedláčková, T. (2022). Opting for Polarizing Emotions: Strategies of Czech Pro-Vaccination Discussants in the Emotionalized Public Sphere and Debate on a Measles Epidemic. *International Journal of Communication*, *16*, 21, 1006-1026.

Wang, X., Yang, L., Yang, J., Wang, P., & Lei, L. (2017). Trait anger and cyberbullying among young adults: A moderated mediation model of moral disengagement and moral identity. *Computers in Human Behavior*, *73*, 519-526.

Wollebæk, D., Karlsen, R., Steen-Johnsen, K., & Enjolras, B. (2019). Anger, fear, and echo chambers: The emotional basis for online behavior. *Social Media+ Society*, *5*(2), 1-14.

Young, M. J., Tiedens, L. Z., Jung, H., & Tsai, M. H. (2011). Mad enough to see the other side: Anger and the search for disconfirming information. *Cognition and Emotion*, *25*(1), 10-21.

Appendices

Appendix A: High Polarization/Social Identity Present

James Thompson @jthompson456 · Oct 18, 2021

Out of curiosity, what do people think about masking during covid?

Amy Angelosi @prnguinkween2 · Oct 18, 2021

Don't listen to the crackpot democrats trying to control us. It's MY RIGHT to not wear a mask and the government should NOT DICTATE what I do with my body #covidhoax #NoMasks

Tom P. @tomtom86 · Oct 18, 2021

Figures a republican would say something like that. Wear a mask moron and save a life! #ProMask #DontBeLame

Dominque @domdotcom3 · Oct 18, 2021

WOW way to make it political. This has nothing to do with politics, this about our right to be who we are. Also, some people can't wear masks due to asthma. #republicanandproud

Jose E. @surfernskater89 · Oct 18, 2021

Wearing a mask prevents Covid from spreading you IDIOTS! Not wearing one puts others at risk, esp. those who can't get vaccinated. #RepublicansAreSelfish

Emilio Sanchez @es1298 · Oct 18, 2021

I thought we had moved beyond discussions about masks

Chantel @daydream5743 · Oct 18, 2021

The CDC doesn't even have the science to back that up--they literally changed there minds a dozen times.

Louisa Louisa @catmom92 · Oct 18, 2021

...You do realize that YOU don't know the facts, right? Covid spreads through the air and is more contagious than flu.#RepublicanLogic?!

Alice the Nomad @twihardpotter19 · Oct 18, 2021

I swaer to gawd if one more #democrat comments on this post im gonna come through this computer and give them common cents #republicansFTW

Carlton Hilton @ch167896 · Oct 18, 2021

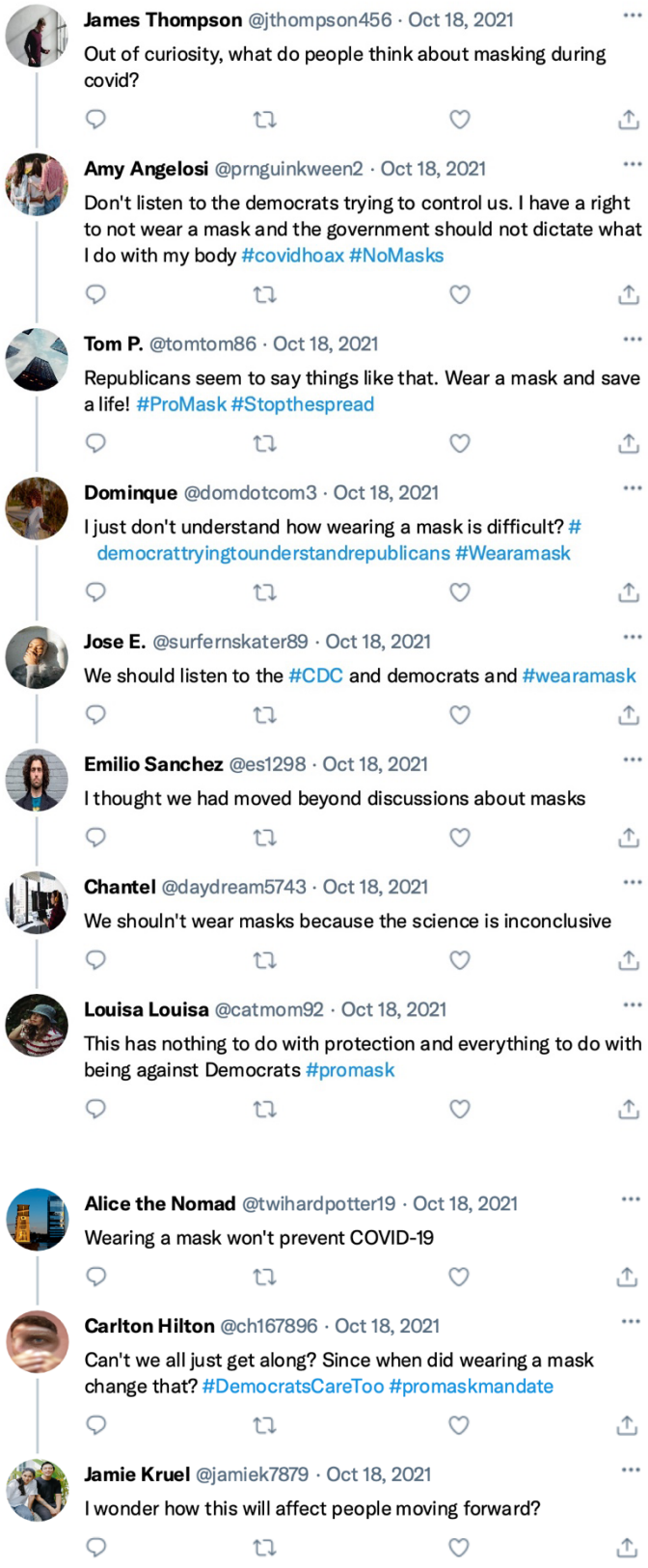
How is wearing a mask taking away your rights? Last time I checked, what you wear isn't even in the Constitution. How about the rights of people YOU'VE make sick by asymtomatic spread?! #DemocratsKnowWhatsGood

Jamie Kruel @jamiek7879 · Oct 18, 2021

I wonder how this will affect people moving forward?

Appendix B: High Polarization/Social Identity Absent

Appendix C: Low Polarization/Social Identity Present



Appendix D: low Polarization/Social Identity Absent

Appendix E: Adapted Moral Outrage Scale

***Directions to give to participants:** Please indicate the extent to which you either agree or disagree with the following statements about masking regarding COVID-19.

***Items to give to participants:**

**Pro-Mask Righteous Anger.

1. Others suffering from a lack of preventative measures concerns me.
2. I am angry about the risks to children and others who can or cannot be vaccinated because people will not wear masks
3. I am angry about how new surges of COVID-19 could affect society because people will not wear masks.

**Pro-Mask Self-Righteous Anger.

1. My own suffering from a lack of preventative measures against COVID-19 concerns me
2. I am angry about the risks to my health because people will not wear masks
3. I am angry about how new COVID-19 surges will affect me because others will not wear masks

**Anti-Mask Righteous Anger.

1. Students' academic life is restricted because of mask wearing
2. I am angry about how mask mandates restrict people's freedom
3. I am angry about how masking has negatively affected businesses, restaurants, gyms, and others

**Anti-Mask Self-Righteous Anger.

1. My academic life has been restricted because of masking
2. I am angry because mask mandates restrict my freedom
3. I am angry about how masking has negatively affected my day-to-day life, including dining out, going out, hanging out with friends and/or family, etc.

***Rated on a Likert scale from 1 to 5, with 1 = Strongly Disagree; 5 = Strongly Agree

Appendix F: Confirmation Bias Scale

***Directions to give to participants:** Please indicate the likelihood that you would share the following information on your social media page.

***Items to give to participants:**

1. The science is inconclusive if wearing a mask prevents the spread of COVID-19.
2. Wearing a mask prevents the spread of COVID-19, similar to covering your nose when sneezing to prevent the spread of infections to others.
3. If you are not sick, you do not need to wear a mask.
4. The science demonstrates that wearing a mask decreases chances of spreading infectious diseases such as COVID-19.
5. The government should not be allowed to mandate wearing masks.
6. We don't wear masks for the flu, so why are they needed for COVID-19?
7. We should all wear masks.
8. Masks make me feel uneasy.
9. I need more information before I trust wearing a mask.
10. There is enough information to justify wearing a mask.

**Rated on a Likert scale from 1 to 5, with 1 = Very Unlikely to Share; 5 = Very Likely to Share

Tables & Figures

Table 1. Means, Standard Deviations, and Correlations Among Variables

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Pro-Mask Righteous Anger	3.65	1.14	--																				
2. Pro-Mask Self-Righteous Anger	2.88	1.24	.771**	--																			
3. Anti-Mask Righteous Anger	2.63	1.18	-.568**	-.486**	--																		
4. Anti-Mask Self-Righteous Anger	2.43	1.23	-.626**	-.525**	.887**	--																	
5. Strength of Political Social Identity	2.52	.89	-.099	-.082	.049	.078	--																
6. Strength of Atittude	2.54	.88	.170**	.129*	-.163**	-.164**	.039	--															
7. Confirmation Bias	2.40	.70	.143**	.193**	.075	.072	.012	.059	--														
8. Selective Attention Bias	3.25	.78	.388**	.395**	-.301**	-.331**	-.150**	.172**	.036	--													
9. Emotional Content Bias	1.70	.52	-.029	-.067	-.051	-.056	.019	.292**	-.080	.027	--												
10. Commitment Bias	1.55	.52	.125*	.067	-.070	-.122*	.015	.473**	.079	.161**	.248**	--											
11. Verbal Reasoning	9.80	6.30	.093	.0116	-.187**	-.144**	.005	.032	-.197**	-.008	.073	.005	--										
12. Trait Anger	1.95	.55	.071	.072	-.005	.004	.103	.011	.010	.038	-.016	-.043	-.113*	--									
13. Extraversion	3.15	.88	-.117*	-.114*	.136*	.126*	.116*	.003	.091	-.162**	-.052	.054	-.197**	.001	--								
14. Agreeableness	3.77	.64	-.008	-.090	.132*	.066	.049	-.077	.061	-.172**	-.037	.004	-.001	-.374**	.242**	--							
15 Conscientousness	3.64	.58	-.042	-.041	.047	.043	.151**	.044	.021	-.016	-.038	.029	-.010	-.083	.183**	.271**	--						
16. Neuroticism	3.28	.74	.222**	.199**	-.140**	-.130*	-.052	.067	-.046	.144**	.157**	.038	.081	.295**	-.350**	-.254**	-.192**	--					
17. Oppeness to Experience	3.41	.66	.211**	.217**	-.045	-.069	-.017	.075	.095	.128*	.087	.032	.010	.001	.101	.060	.016	-.033	--				
18. Ingroup Loyalty	3.48	.85	-.383**	-.287**	.475**	.441**	.150**	-.148**	.091	-.239**	-.036	-.062	-.148**	.042	.219**	.245**	.149**	-.178**	-.018	--			
19. Proneness to Confirmation Bias	33.32	5.20	-.061	-.063	.180**	.207**	.135*	-.062	.105*	-.044	-.128*	-.059	-.156**	.301**	.166**	.045	.161**	-.065	.060	.324**	--		
20. Proximity to COVID-19	2.34	1.05	-.125*	-.147**	.010	.093	-.065	.010	-.098	-.150**	.005	-.046	.102	-.000	.065	.071	.011	.080	-.078	.005	.002	--	
21. Pre-Stance Attitude on Masking	2.43	.69	.682**	.584**	-.540**	-.607**	-.149**	.183**	-.028	.376**	-.032	.118*	.082	-.001	-.110*	-.069	-.021	.131*	.195**	-.404**	-.109*	.114	--

Note. $n = 357$; $p < .05$*; $p < .001$**

Table 2. Outcomes Assessed with Explicit Reference to Social Identity in Twitter Feeds.

	Social Identity Absent		Social Identity Present				
Outcome Assessed	*M*	*SD*	*M*	*SD*	*t*(355)	*p*	*Cohen's D*
Pro-Mask Righteous Anger	3.56	1.20	3.73	1.07	-1.44	.150	1.14
Pro-Mask Self-Righteous Ang	2.75	1.20	3.02	1.27	-2.03	.044*	1.24
Anti-Mask Righteous Anger	2.62	1.19	2.65	1.17	-.26	.790	1.18
Anti-Mask Self-Righteous An	2.39	1.23	2.48	1.24	-.66	.510	1.23
Strength of Attitude	2.55	.92	2.56	.84	-.383	.702	.88
Confirmation Bias	2.36	.66	2.43	.73	-.972	.332	.70
Selective Attention Bias	3.21	.77	3.28	.79	-.85	.394	.78
Emotional Content Bias	1.73	.54	1.66	.51	1.30	.195	.52
Commitment Bias	1.57	.51	1.53	.54	.70	.484	.52

Note. $p < .05$*; $p < .001$**

Table 3. Outcomes Assessed with Content Polarization in Twitter Feeds.

	Low Polarization		High Polarization				
Outcome Assessed	*M*	*SD*	*M*	*SD*	*t*(355)	*p*	*Cohen's D*
Pro-Mask Righteous Anger	3.70	1.14	3.59	1.14	.89	.373	1.14
Pro-Mask Self-Righteous Anger	2.98	1.25	2.79	1.23	1.45	.149	1.24
Anti-Mask Righteous Anger	2.66	1.19	2.61	1.17	.36	.708	1.18
Anti-Mask Self-Righteous Anger	2.45	1.21	2.41	1.26	.31	.756	1.23
Strength of Attitude	2.46	.88	2.63	.87	-1.87	.062*	.87
Confirmation Bias	2.48	.63	2.31	.76	2.30	.022*	.69
Selective Attention Bias	3.22	.70	3.27	.85	-.547	.585	.78
Emotional Content Bias	1.59	.44	1.81	.58	-4.07	.001**	.51
Commitment Bias	1.53	.49	1.57	.56	-.72	.475	.52

Note. $p < .05$*; $p < .001$**

Table 4. ANCOVAs Assessing the Relationships of Social Identity and Content Polarization on Different Types of Self-Righteous Anger

	Pro-Mask Self-Righteous Anger			Anti-Mask Self-Righteous Anger		
Variable	*F*	*p*	*η2*	*F*	*p*	*η2*
Corrected Model	31.03	.001**	.384	50.36	.001**	.418
Intercept	2.39	.123	.007	96.95	.001**	.216
In-Group Loyalty	--	--	--	27.13	.001**	.072
Neuroticism	9.58	.002*	.026	--	--	--
Openness to Experience	5.98	.015*	.017	--	--	--
Proximity to COVID-19	4.17	.041*	.012	--	--	--
Pre-Stance Attitude on Masking	149.05	.001**	.299	133.71	.001**	.276
Polarization (i.e., High, Low)	1.45	.229	.004	.35	.557	.000
Social Identity (i.e., Absent, Present)	2.39	.123	.007	1.88	.171	.005
Polarization x Social Identity	.01	.915	.000	.11	.742	.000

Note. Only significant relationships with covariates are represented in this table. $p < .05$*; $p < .001$**

Table 5. ANCOVAs Assessing the Relationships of Social Identity and Content Polarization on Strength of Attitude.

	Strength of Attitude		
Variable	*F*	*p*	*η2*
Corrected Model	4.14	.003*	.045
Intercept	135.31	.001**	.278
Pre-Stance Attitude on Masking	12.80	.001**	.035
Polarization (i.e., High, Low)	3.99	.046*	.011
Social Identity (i.e., Absent, Present)	.07	.795	.000
Polarization x Social Identity	.11	.774	.000

Note. Only significant relationships with covariates are represented in this table. $p < .05$*; $p < .001$**

Table 6. ANCOVAs Assessing the Relationships of Social Identity and Content Polarization on Cognitive Biases.

Variable	Confirmation Bias			Selective Attention Bias			Emotional Content Bias			Commitment Bias		
	F	*p*	*n2*	*F*	*p*	*n2*	*F*	*p*	*n2*	*F*	*p*	*n2*
Corrected Model	5.81	.001**	.062	13.14	.001**	.158	7.54	.001**	.079	1.57	.182	.018
Intercept	1536.78	.001**	.814	199.13	.001**	.362	112.73	.001**	.243	170.96	.001**	.327
Verbal Reasoning	12.36	.001**	.040	--	--	--	--	--	--	--	--	--
Pre-Stance Attitude on Masking	--	--	--	54.88	.001**	.135	--	--	--	5.19	.023	.015
Proximity to COVID-19	--	--	--	5.056	.025	.014	--	--	--	--	--	--
Neuroticism	--	--	--	--	--	--	11.31	.001**	.031	--	--	--
Polarization (i.e., High, Low)	6.16	.014*	.017	.68	.409	.002	17.68	.001**	.048	.60	.439	.002
Social Identity (i.e., Absent, Present)	.90	.343	.003	.31	.578	.001	2.55	.111	.007	.61	.435	.002
Polarization x Social Identity	1.48	.224	.004	.96	.327	.003	.35	.554	.001	.02	.896	.000

Note. Only significant relationships with covariates are represented in this table. $p < .05$*; $p < .001$**

Table 7. Continuous Variables Moderation Analyses using the PROCESS Macro.

Outcome Assessed		Independent Variables and Covariates	*b*	*SE*	*t*	*p*	LLCI	ULCI
Pro-Mask Self-Righteous Anger	*Model 1*	Neuroticism	.23	.07	3.12	.002*	.08	.37
		Openness to Experience	.20	.08	2.49	.013*	.04	.36
		Proximity to COVID-19	-.10	.05	-2.05	.041*	-.20	-.00
		Pre-Stance Attitude on Masking	.96	.08	12.06	.001**	.80	1.12
		Social Identity	.23	.32	.73	.468	-.39	.85
		Strength of Social Identification	.04	.18	.21	.831	-.32	.40
		Social Identity * Strength of Social Identification	-.03	.12	-.22	.825	-.26	.21
Anti-Mask Self-Righteous Anger	*Model 2*	Ingroup Loyalty	.34	.07	5.21	.001**	.21	.47
		Pre-Stance Attitude on Masking	-.93	.08	-11.58	.001**	-1.08	-.77
		Social Identity	.04	.30	.14	.885	-.55	.64
		Strength of Social Identification	-.10	.18	-.58	.564	-.45	.24
		Social Identity * Strength of Social Identification	.04	.11	.33	.743	-.19	.26
Strength of Attitude	*Model 3*	Pre-Stance Attitude on Masking	.24	.07	3.64	.002*	.11	.38
		Social Identity	.21	.28	.76	.446	-.33	.75
		Strength of Social Identification	.18	.16	1.10	.273	-.14	.49
		Social Identity * Strength of Social Identification	-.07	.10	-.72	.470	-.28	.13

Confirmation Bias	*Model 4*	Verbal Reasoning	-.02	.01	-3.77	.001**	-.03	-.01
		Social Identity	.12	.22	.56	.574	-.31	.55
		Strength of Social Identification	.04	.13	.33	.741	-.21	.29
		Social Identity * Strength of Social Identification	-.02	.08	-.26	.792	-.18	.14
Selective Attention	*Model 5*	Proximity to COVID-19	-.09	.04	-2.39	.018*	-.16	-.02
		Pre-Stance Attitude on Masking	.39	.06	6.93	.001**	.38	.50
		Social Identity	.07	.23	.32	.747	-.38	.52
		Strength of Social Identification	-.07	.13	-.56	.576	-.34	.19
		Social Identity * Strength of Social Identification	-.01	.09	-.14	.890	-.18	.16
Emotional Content Bias	*Model 6*	Neuroticism	.12	.04	3.27	.001**	.05	.19
		Polarization	.38	.16	2.33	.020*	.06	.70
		Strength of Social Identification	.12	.10	1.24	.216	-.07	.31
		Polarization * Strength of Social Identification	-.06	.06	-.93	.353	-.18	.06

Note. Only significant relationships with covariates are represented in this table. $p < .05$*; $p < .001$**

Table 8. Moderated Mediation Analyses using the PROCESS Macro

Outcome Assessed		Variables	Direct Effect	Indirect Effect	Moderator	*SE*	LLCI	ULCI
Strength of Attitude	*Model 1*	Polarization*	.19*		--	.09	.01	.37
		Pro-Mask Righteous Anger	--	-.04	--	.04	-.13	.01
		Pro-Mask Self-Righteous Anger	--	-.00	--	.01	-.02	.02
		Polarization * Social Identity	--	--	.00	.03	-.06	.05
	Model 2	Pre-Stance Attitude on Masking**	.18	--	--	.08	.01	.34
		Polarization	.18	--	--	.09	-.00	.36
		Anti-Mask Righteous Anger	--	.01	--	.02	-.03	.05
		Anti-Mask Self-Righteous Anger	--	-.00	--	.01	-.02	.01
		Polarization * Social Identity	--	--	-.00	.02	-.04	.04
Confirmation Bias	*Model 3*	Verbal Reasoning**	-.02	--	--	.01	-.03	-.01
		Polarization*	-.16	--	--	.07	-.30	-.02
		Pro-Mask Righteous Anger	--	-.01	--	.02	-.05	.03
		Pro-Mask Self-Righteous Anger	--	-.00	--	.01	-.03	.02
		Polarization * Social Identity	--	--	.03	.03	-.00	.09
	Model 4	Verbal Reasoning**	-.02	--	--	.01	-.03	-.01
		Polarization*	-.18	--	--	.07	-.32	-.04
		Anti-Mask Righteous Anger	--	.00	--	.01	-.03	.03
		Anti-Mask Self-Righteous Anger	--	.00	--	.01	-.01	.02
		Polarization * Social Identity	--	--	-.00	.02	-.04	.03
Selective Attention Bias	*Model 5*	Pre-Stance Attitude on Masking*	.19	--	--	.08	.05	.34
		Agreeableness*	-.18	--	--	.06	-.29	-.06
		Social Identity	-.01	--	--	.07	-.16	.13
		Pro-Mask Righteous Anger	--	.01	--	.01	-.01	.05
		Pro-Mask Self-Righteous Anger	--	.01	--	.01	-.02	.04
		Social Identity * Strength of Social Identification	--	--	-.01	.01	-.04	.01

	Model 6	Pre-Stance Attitude on Masking**	.30	--	--	.07	.16	.44
		Agreeableness*	-.17	--	--	.06	-.29	-.05
		Social Identity	.04	--	--	.08	-.11	.19
		Anti-Mask Righteous Anger	--	.00	--	.01	-.02	.02
		Anti-Mask Self-Righteous Anger	--	-.01	--	.01	-.03	.01
		Social Identity * Strength of Social Identification	--	--	-.01	.02	-.05	.01
Emotional Content Bias	*Model 7*	Neuroticism**	.13	--	--	.13	.05	.20
		Polarization**	.22	--	--	.06	.11	.33
		Pro-Mask Righteous Anger	--	-.00	--	.01	-.02	.01
		Pro-Mask Self-Righteous Anger	--	-.00	--	.01	-.01	.01
		Polarization * Strength of Social Identification	--	--	-.00	.01	-.02	.01
	Model 8	Neuroticism*	.11	--	--	.04	.04	.19
		Polarization**	.23	--	--	.05	.12	.33
		Anti-Mask Righteous Anger	--	-.00	--	.01	-.02	.02
		Anti-Mask Self-Righteous Anger	--	-.00	--	.00	-.01	.01
		Polarization * Strength of Social Identification	--	--	-.00	.01	-.02	.01
Commitment Bias	*Model 9*	Polarization	.04	--	--	.06	-.07	.15
		Pro-Mask Righteous Anger	--	-.03	--	.02	-.08	.00
		Pro-Mask Self-Righteous Anger	--	.00	--	.01	-.01	.02
		Polarization * Social Identity	--	--	-.01	.02	-.05	.02
	Model 10	Polarization	.04	--	--	.05	-.07	.15
		Anti-Mask Righteous Anger	--	-.00	--	.02	-.05	.04
		Anti-Mask Self-Righteous Anger	--	-.01	--	.01	-.04	.01
		Polarization * Social Identity	--	--	.01	.03	-.05	.08

Note. Only significant relationships with covariates are represented in this table. $p < .05$*; $p < .001$**