Theoretical and Methodological Advancements in Psychological Reactance Theory: Examining Interactions between Individuals, Messages, and Social Contexts

By

Lauren A. Kriss

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Journalism and Mass Communication)

at the

UNIVERSITY OF WISCONSIN-MADISON

2024

Date of final oral examination: 08/05/2024

This dissertation is approved by the following members of the Final Oral Committee:

    Christopher N. Cascio, Associate Professor, Journalism and Mass Communication

    Sijia Yang, Associate Professor, Journalism and Mass Communication

    Rasmus Birn, Associate Professor, Wisconsin Psychiatric Institute and Clinics

    Dhavan Shah, Maier-Bascom Professor, Journalism and Mass Communication

    Brian L. Quick, Professor, Department of Communication, School of Medicine, University of Illinois at Urbana-Champaign

Table of Contents

Acknowledgments

I am the least self-made person, and for that I am so grateful. I hope I can always earn and repay the support so many people have given me on my PhD journey. Thank you to my advisor Dr. Christopher N. Cascio for his support in the writing and revision of this dissertation and for the opportunities I had in the Communication, Brain and Behavior lab, including the data collection opportunities related to cannabis warning labels and public service announcements. Thank you to the committee for investing their time in evaluating this dissertation. Thank you to committee member Dr. Rasmus Birn for being an excellent instructor in functional magnetic resonance imaging. Thank you to Dr. Dhavan Shah for an excellent seminar in academic publishing that helped prepare me for the course ahead.

Thank you to Dr. Sijia Yang for serving on my committee and for leading the population survey related to cannabis warning labels, which study 3 is a companion to. Thank you to my colleagues in the CAMER lab who worked on the warning label project, including Lynne M. Cotter and Linqi Lu. Thank you to Dr. Jiaying Liu for supporting the study of cannabis warning labels as well.

Thank you to Dr. Brian Quick for serving on the committee and for providing the opportunity to collect data related to vaccine mandates. Also, thank you for guiding me through my master's degree during a particularly hard time and for always taking my ideas seriously. You are a best friend for life.

I also want to acknowledge my main co-collaborator during my time at the University of Wisconsin, Dr. Matt Minich, for all the brainstorming sessions, brains scanned, and crises averted. I couldn't have finished without him. I also want to acknowledge my other friends in the

supportive of my academic pursuits. Thank you, Dr. Ruth Rechis, Martha Vieco-Garcia, Haley Gardiner, Mayra Aquino, Lauren McDonald, Jasmine Dailey and Terrance Adams.

I owe so much to my personal support system outside of academia. I love my hometown friends for knowing me for almost two decades and never caring about my job. Our group chat has always been a lifeline. Laurel Brown, Alyssa Cepeda, Alina Tang, and Holly Hundsdorfer, I love you so much. Thank you to my best friends from Texas Tech, Angus Robertson, Jake Noltensmeyer, Amanda Benge and Jacob Howle, for always encouraging me and checking in.

Everything I am, I am because of my family. Thank you to my Grammy, Nicole Kriss, for always checking in, being understanding of when I am busy and being a cheerleader. Thank you to my aunt, Traci Kriss, for being the embodiment of 'work, hard, play hard.' I trust you have the proper bubbly ready to celebrate. Thank you to my twin brother, Carter Kriss, for loving a good argument. Most of all, I owe gratitude to my parents, Kevin and Kellie Kriss. The foundation you built is the reason for all the good things in my life. My dad has helped me move apartments and states many times. Thank you for always being there for me, both when I need emotional support and when I just *need help*. I also want to thank my mom because she worked very hard to give me opportunities she never had, and she has always encouraged me to push boundaries. For both my parents, I hope I can be as strong and loving as you.

To my family and friends: Thank you and I love you more than I could ever express.

Abstract

Theoretical and Methodological Advancements in Psychological Reactance Theory: Examining Interactions between Individuals, Messages, and Social Contexts

Lauren A. Kriss

Psychological reactance theory offers important insights into how individuals may respond to threats to freedom, including when individuals will resist or acquiesce. Most contemporary research has focused on operationalizing and measuring reactance alongside studying message features that incite or diminish reactance. There is a need to understand how individual personality differences, social context and messages interact with threats to freedom to predict motivation. Furthermore, an earlier focus on inciting reactance means the field knows less about how messages motivate reactance in more realistic contexts. This dissertation seeks to improve understanding of motivation in response to freedom threats across three studies. Study 1 showed that two trait-level characteristics, reactance proneness and locus of control, influenced responses to messages about COVID-19 vaccine mandates on college campuses. Study 2 used focus groups to show that differences in the social dynamics of groups, specifically how much they know about other participants' behavior, influenced the responses to public service announcements regarding the risks of driving under the influence of cannabis. Finally, study 3 searched for reactance responses to cannabis warning labels using functional magnetic resonance imaging and found no evidence of reactance. Taken together, this dissertation showed that reactance is influenced by individual personalities, specific messages, and group dynamics. The insights from this dissertation suggest theoretical and methodological improvements that can be made to psychological reactance theory.

Introduction and Overview

The theories that form the vast field of persuasion psychology exist in roughly three categories a) behavior change theories, b) theories of information processing and c) theories of message effects (Cappella, 2006). A 2006 special issue of *Journal of Communication* (Cappella & Rimer, 2006) attempted to integrate behavior change theories, information processing theories, and message effects theories to advance understanding of mechanistic causality. Cappella (2006) argued that if the field of persuasive message design does not know why and how a message feature influences a belief, then "we are at a loss for connecting findings with [a] message feature to other findings on different features that employ the same mechanism," (Cappella, S267). By linking behavior change, information processing and message effects, it is possible to discover when the relationships between individuals, messages and social contexts hold to engage in evidence-based persuasion.

Lewin (1939) described the "psychological field," as "the person in the situation" (p. 5) such that the prediction line of what an individual will do cuts across the field based on motivation (Fiske & Taylor, 1991). Behavior change theories offer potential routes of persuasion based on beliefs, whereas information processing theories exist to help identify *mechanisms* to belief change based on individuals context and work with theories of message effects to craft communication (Cappella, 2006). A core function of processing theories is to understand motivation.

The most famous processing theories are the elaboration likelihood model (ELM) (Petty & Cacioppo, 1986), the Limited Capacity Model of Motivated Message Processing (LC4MP) model (Lang, 2000) and the heuristic-systematic model of information processing (Chaiken, 1980). Another well-known theory that shares methodologies and overlaps theoretically with

other information processing theories is Brehm's (1966) psychological reactance theory (PRT). Fundamentally, the goal of psychological reactance theory is to predict when the motivation to restore freedom will lead individuals to resist or acquiesce (Brehm & Brehm, 1981) depending on the individual and the context of the freedom threat.

Historical Advancements in Reactance Theory

In their 50-year review of PRT, Rosenberg and Siegel (2018) review the transdisciplinary origins and history of PRT and urge new scholars to harken back to the focus on understanding motivation. In their history, there are five waves of PRT research, and the literature is currently in the fifth wave: Wave 1: Theory proposal and testing, Wave 2: Contributions from clinical psychology, Wave 3: Contributions from communication research, Wave 4: Measurement of reactance, and Wave 5: Return to motivation.

The first wave is when the initial assumptions and theoretical proposition were tested, mostly with small-scale laboratory experiments. The second wave came out of clinical psychology and added the concept of trait reactance to the literature and the formal theory (Brehm & Brehm, 1081), though controversy over trait reactance remains (Rosenberg & Siegel, 2018). From a clinical psychology perspective, trait reactance was viewed as a negative moderator of clinical treatment, such that individuals who were more reactant would experience less benefit from therapy and were less likely to adhere to treatment plans (for review see Shoham et al., 2004). Ultimately, this thinking also led to research on overcoming reactance in clinical psychology and designing treatments based on trait level reactance (Rosenberg & Siegel, 2018). In many ways, this idea of tailoring messages based on trait reactance, also known as audience segmentation, is what persuasion studies often borrow from the clinical psychology perspective both to study the associations between trait reactance and health risk behavior (e.g.

Miller & Quick, 2010) and to predict reactance to health communication materials (e.g. LaVoie et al., 2017; Reynolds-Tylus, 2019).

Wave three, contributions from communication research, happened concurrently with contributions from clinical psychology, according to Rosenberg and Siegel (2018). This wave, which in many ways still continues, focused on message features that could both incite and diminish reactance responses to persuasive messages, which are conceived of as freedom threats (Quick et al., 2013; Reynolds-Tylus, 2019; Rosenberg & Siegel, 2018). Brehm and Brehm (1981) did theorize about some message features, but it was not a predominant focus until communication researchers took up the framework.

In wave four social psychologists, many of them communication researchers, broke away from the original theoretical texts (Brehm, 1966; Brehm & Brehm, 1981) to argue that reactance was not immeasurable as the original theorists stated (Ratcliff, 2021; Rosenberg & Siegel, 2018). The first study in this vein, considered seminal, was Dillard and Shen's (2005) test of competing structural models. The intertwined model proposed by Dillard and Shen (2005), which models reactance as a latent factor consisting of anger and negative cognition, predominates still today and has held up across meta-analysis and many reviews (Rains, 2013; Ratcliff, 2021; Reynolds-Tylus, 2019).

Dissertation Overview: Return to Motivation

Rosenberg and Siegel (2018) called for a fifth wave: Return to motivation. They called for three new directions to advance research in reactance by examining a) individual factors, specifically states of mind, b) catalysts of reactance outside of direct freedom threats, particularly

implicit threats, and c) new outcomes of reactance. This dissertation seeks to answer this call across three studies.

Study 1 examines two famous personality variables thought to influence state reactance: trait reactance and locus of control. A core function of processing theories is to explain how individual differences drive motivation in various contexts, and PRT should be no exception. Further, a better understanding of the ways that personality variables predict behavior in response to persuasion would allow for strategic audience segmentation when designing persuasive messages (Quick et al., 2013). The studies of individual differences in reactance are limited, and the field needs new traits to investigate (Rosenberg & Siegel, 2018). Additionally, the previous evidence needs to be updated with better methodologies. Although trait reactance is controversial, Study 1 does show the predictive utility of the construct. Study 1 also returns locus of control to its historical roots in studies of activism as students respond to messages about policy mandates. Further, study 1 explores interactions between individual differences and properties of freedom threats to predict freedom threat and attitudes, with a focus on indirect threats.

Study 2 investigates the social dynamics that influence reactance to persuasive messages by coding focus group conversations for negative cognitions and anger. This study helps uncover the potential influence of implicit factors that influence opinions, including knowledge of behavioral norms in a group. More research using qualitative techniques would be valuable to triangulate understanding between self-reported measures and how reactance unfolds in interpersonal discussions. In particular, Study 2 uses focus groups to begin exploring the ways that mediated communication, discussion and reactance interact depending on differences in group make up. The original theorizing on social power and reactance was limited to dyads and

interpersonal attempts at influence (Brehm & Brehm, 1981), but Study 2 attempts to simulate group dynamics in the context of co-viewing persuasive messages. The goal of Study 2 is to begin providing insight into how persuasion is appraised in an era where a large portion of media content is consumed through networked communication in social groups in the offline and online world.

Finally, Study 3 investigates neural activity in regions associated with behavior change in response to persuasive messages (ventral medial prefrontal cortex (VMPFC); Bartra et al., 2013; Falk et al. 2010a, b; 2012) and neural activity in regions associated with anger (amygdala) and negative affect (dorsal anterior cingulate (dACC) and anterior insula (AI)) (Barrett & Bliss-Moreau, 2009; Richard et al., 2023) that may underlie the emotional component of reactance (Dillard & Shen, 2005, Ratcliff, 2021). Study 3 also tested the ability of anger and negative affect regions to predict self-reported anger, one of the primary components of reactance (Quick et al., 2013; Ratcliff, 2021). The primary benefit of utilizing a neuroscience approach is to establish that the experience of reactance is not an artifact of self-reported measures, particularly as the emotional component of reactance is not well-suited to self-report measures.

Across all three studies, there is a consistent effort to use real-world messages. Due to the recent focus of communication researchers on operationalizing and measuring reactance itself (Rains, 2013), much of the research attempts to incite reactance directly using contrived messages. There has been less of a focus on seeing what messages in everyday persuasion and health communication incite reactance. Further, although the message effects literature in PRT is extensive, it relies heavily on tightly controlled experimental messages lacking in external validity.

The remainder of the introduction chapter seeks to provide the necessary theoretical and methodological overview of PRT, including constructs, assumptions and propositions of PRT. Further, there is a preview of the unresolved questions, shortcomings and controversies in the reactance literature that this dissertation, at least in part, attempts to add evidence for.

Overview of Psychological Reactance Theory

PRT seeks to offer explanations for how individuals will respond to freedom threats and seeks to predict who, in what situations, will be motivated to restore their freedom (Brehm & Brehm, 1981). The implications of PRT are vast, including interpersonal and mass-mediated social influence, clinical psychology, policy mandates, even romantic encounters and shopping (Brehm & Brehm, 1981). The core principles and constructs of PRT explain why the theory possesses so much explanatory power. Fundamentally, the magnitude of reactance is determined by the value of the freedom threatened and by the nature of the threat to freedom.

Key Constructs in Reactance Theory

Freedom

In the framework of PRT, the first principle is that freedom is a subjective, perceived construct. Specifically, Brehm and Brehm (1981) stated that "freedom is an expectancy and can be held with more or less certainty," (p. 5) and that reactance can only be expected "to the extent that the individual believes he or she has a freedom or control over an outcome," (p. 5). Importantly, this is why PRT can predict both backlash, compliance and even the complete surrender of individual freedom. Reactance is also a function of the perceived importance of any freedom to an individual within PRT (Brehm & Brehm, 1981).

Freedom Threat

The principles of PRT that determine the magnitude of reactance from the side of the freedom threat are the number of freedoms threatened and the direct or indirect nature of the threat. Therefore, freedom threats cover a wide variety of social phenomena, from prices to persuasion attempts to brutally enforced repression of individual liberties. Brehm and Brehm (1981) stated, when describing threats to freedom that "the individual's control over each potential outcome is threatened or reduced by an increase in the perceived difficulty of attaining an outcome," (p. 3).

The research testing the properties of freedom threats and their subsequent effects on reactance is limited. Most recently, Kriss et al. (2022) experimentally manipulated a vaccine mandate message to either include reference to sanctions, or not, and tested the difference between direct and indirect threats. Direction of threat was manipulated by whether participants read a message about a vaccine mandate at their own college or another college instead. Consistent with PRT, the greater magnitude threat generated more reactance. Surprisingly, an interaction appeared such that the most reactance was generated by an indirect threat with sanction, leading the authors to speculate about the role of uncertainty.

Freedom threat and reactance overlap in the PRT literature, particularly in research that uses self-report measurements (Ratcliff, 2021). Specifically, "freedom threat is sometimes treated as an antecedent and other times as a measure of reactance itself," (p. 7) which means that there are studies which combine freedom threat and reactance into a single measure and studies that don't measure freedom threat at all (Ratcliff, 2021). Subjective measures of freedom threat, such as the original from Dillard and Shen (2005), rely on an individual evaluating how much a message is trying to manipulate them or threaten their freedom to choose. However, those same sentiments could plausibly be argued as reactance itself.

Psychological Reactance

Reactance is "the motivational state that is hypothesized to occur when a freedom is eliminated or threatened with elimination," (Brehm & Brehm, 1981, p. 37; Quick et al., 2013). As a motivational state, reactance "is assumed to have energizing and behavior-directing properties," (Brehm & Brehm, 198, p. 98).

Brehm and Brehm never stated what constitutes reactance itself and how it is different from other types of motivations, focusing instead on antecedents and consequences. (Brehm, 1966; Brehm & Brehm, 1981; Dillard & Shen, 2005). Dillard and Shen noted the conspicuous absence of details about the nature of reactance:

> "Apart from a brief mention of the possibility that individuals "may be aware of hostile and aggressive feelings" (Brehm, 1966, p. 9), if the level of reactance arousal is high, the nature of reactance itself remains remarkably underdetermined. Brehm's apparent reluctance to provide greater conceptual explication of the principal mechanism of the theory may arise from his belief concerning the potential for measurement of it. According to Brehm and Brehm (1981), "reactance has the status of an intervening, hypothetical variable ... We cannot measure reactance directly, but hypothesizing its existence allows us to predict a variety of behavioral effects" (p. 37)," (Dillard & Shen, 2005, p. 146).

Although the original theorists of PRT held that reactance cannot be observed, communication scholars have pushed a research agenda to operationalize and measure reactance (Quick et al., 2013; Ratcliff, 2021; Rosenberg & Siegel, 2018). This work was formalized by Dillard and Shen (2005) in a seminal theory piece. Ultimately, Dillard and Shen posited justification for four distinct models of what reactance is, a) cognition, b) affect, c) cognition and

affect or d) an amalgamation of anger and negative cognition. A series of models tested originally, then meta-analytically supported a decade later by Rains (2013), aligns with the perspective that anger and negative cognitions are intertwined:

> "…cognition and affect are intertwined. In fact, they are intertwined to such a degree that their effects on persuasion cannot be disentangled. Such a view is most compatible with a conception of motivation as an alloy of its components, rather than a simple sum of distinct elements," (Dillard & Shen, 2005, p. 147)."

Perhaps as a symptom of the theory's original focus on antecedents and consequences, the theoretical work on reactance and freedom restoration is also muddled by 'conceptual overlap' between indicators of reactance cognitive responses and freedom restoration behaviors (Ratcliff, 2021). For the sake of clarity, this piece relies on the conceptual categories laid out by Ratcliff (2021) which separates cognitive and affective indicators of reactance from restoration outcomes, particularly attitudes, intention, and behaviors.

Cognitive indicators of reactance in the literature include negative cognitions, counterarguing, perceived argument quality, message minimization or derogation, message evaluation and source evaluation (Ratcliff, 2021). The PRT literature has given much less attention to affective indicators of reactance, instead focusing almost exclusively on anger (Ratcliff, 2021). The vast majority of research uses the self-reported anger scale from Dillard and Shen (2005). Most PRT research does not code thoughts generated by participants for anger (Ratcliff, 2021), as Dillard and Shen (2005) did originally. However, it is certainly possible and may aid qualitative investigations into reactance, as shown in Study 2.

There are still concerns about the current measurement of reactance. Communication scholars have called for greater efforts to triangulate reactance across measures, with a particular focus on biological measurement (Rains, 2013; Ratcliff, 2021). Therefore, neural indications of anger and negative affect, which may underlie reactance are considered in Study 3.

Freedom Restoration

Brehm and Brehm (1981) held that behaviors undertaken to restore freedom were the "direct manifestation," (p. 4) of reactance. Brehm and Brehm also theorized about indirect attempts to restore freedom. Specifically, freedom can be restored through a) increasing liking to the alterative position of the freedom threat, b) engaging in the behavior directly, c) engaging in a substitute for the behavior and d) encouraging another individual to engage the behavior (Brehm, 1966; Ratcliff, 2021).

Attitude change is the most relevant to persuasion. In the original theorizing, Brehm and Brehm (1981) followed the definition that attitudes are evaluative and communicate underlying values. Further, they distinguish between two 'attitudinal freedoms': the freedom to take a position and the freedom to not take a position. In this view, changing someone's attitude and attempting to make an individual take a position can both be reactance-inducing. Very little research has been done around the freedom to take no position, but recently Quick et al. (2024) speculated that in the context of opt-out organ donation, messages framed as supportive or in opposition to opt-out organ donation incited the same amount of freedom threat. They speculated that the freedom to take no position was valued highly on an issue that most Americans do not consistently follow.

A huge volume of literature, including the original intertwined model from Dillard and Shen (2005) view attitude change in the opposite direction of the advocated position, as a form of freedom restoration (Ratcliff, 2021; Reynolds-Tylus, 2019). Controversially, some scholars even treat failures to change attitudes as evidence of reactance (Ratcliff, 2021). Now, with all the constructs of PRT explicated, the process theory of reactance will be explained.

Theorizing Reactance as a Process

Like the contemporary discussion around the nature of reactance, the contemporary re-imagining of reactance as a process model that can be observed was theorized and tested first by Dillard and Shen (2005).

Dillard and Shen (2005) reviewed the evidence for four models, and ultimately tested all four on the same data sets to see which best fit the data. See Figure 1. In all models, "antecedents of reactance" represent the "strength of the threat to freedom" and the personality trait thought to encapsulate a "trans-situational propensity to experience reactance," (Dillard & Shen, 2005, p. 149). Although Dillard and Shen speculated about interaction effects from both the threat and trait reactance in their initial "antecedents" variable, most scholars now model freedom threat as the only antecedent, and trait reactance research is studied infrequently (Rains, 2013; Rosenberg & Siegel, 2018). The controversies surrounding the role of trait reactance, and the way reactance scholars should think about reactance proneness are evaluated and tested in Study 1.

The main differences across the models that Dillard and Shen (2005) and later Rains (2013) were looking to resolve were tied to the nature of reactance itself, as discussed previously. Ultimately, the model that had anger and negative cognition intertwined has won out (Dillard & Shen, 2005; Rains, 2013; Ratcliff, 2021).

The other contribution of communication scholars was theorizing reactance as a process model in which the effect of threats on behavior is mediated by reactance and subsequent attitude change (Rosenberg & Siegel, 2018). Although Brehm and Brehm (1981) explicitly theorized reactance as a mediating state between threats and attempts to restore freedom, it wasn't until the introduction of structural equation modelling to communication science that the mediation paths were tested explicitly and supported through subsequent studies and meta-analyses (Dillard & Shen, 2005; Rains, 2013; Rosenberg & Siegel, 2018). An innovation of communication theory, Dillard and Shen (2005) in particular, is the use of reasoned action models (see Fishbein and Ajzen, 2010) to include the dual mediation of reactance and attitude change to predict behavior.

Although the model has been supported across studies using standard criteria for structural equation modelling, studies of reactance are still limited by a reliance on self-reported data often collected at the same time point (Rains, 2013; Ratcliff, 2021). A proper analysis of a communication process would go beyond variance-based methods (i.e. structural equation modelling) to examine time sequence and how different versions of a process can unfold (Poole, 2013). Poole (2013) articulated a hybrid approach which uses both variance-based statistical techniques and what Poole termed "process approaches." Reactance research has already done this in many respects but could be strengthened by longitudinal designs and by studies that only manipulate one mediator at a time (Lemmer & Gollwitzer, 2017).

Message Features to Induce and Diminish Reactance

Communication researchers, particularly those interested in health behavior change, took up reactance theory in order to predict backlash effects, but added to the theory by attempting to examine the moderating role of message features in creating and diminishing freedom threats (Reynolds-Tylus, 2019; Rosenberg & Siegel, 2018).

A wide variety of message features have been tested in the PRT literature. In the most comprehensive review of message features and PRT, Reynolds-Tylus (2019) identified freedom-threatening language as the only real driver of increased threat. But, Quick et al. (2013) has suggested that message variables in reactance are not tied strongly enough to academic theories about language. Further, message effects studies suffer from the paradoxical issues associated with effects-based message variables, as articulated by O'Keefe (2003). Reynolds-Tylus (2019) stressed the need for replication studies and noted that the message features studies should increase ties to theory, as an echo of early comments made by Quick et al. (2013).

Figure 1. Process model of reactance
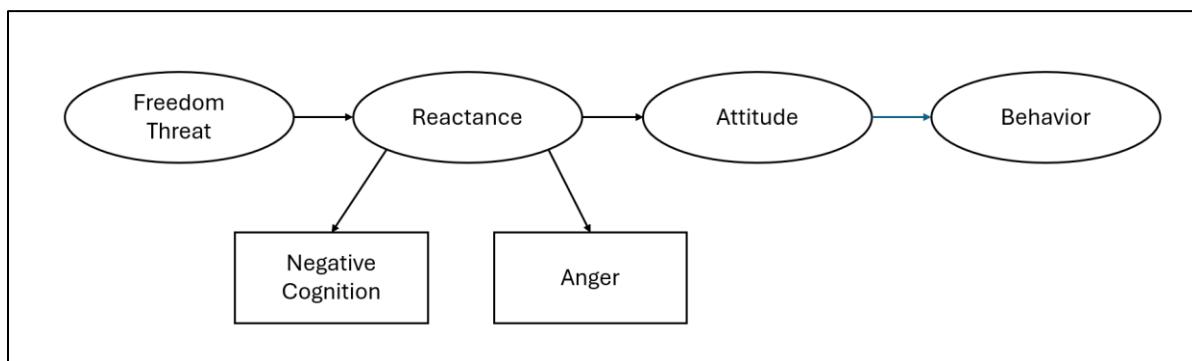
PRT and Communication Research

The ability to reliably test message features to predict successful persuasion, failed persuasion and backlash effects is the primary utility of reactance theory for communication researchers. This dissertation attempts to address previously noted limitations in reactance scholarship. Ultimately, a more refined understanding of the social psychological processes that drive reactance, triangulated across methods, is essential to advance scholarship in anticipating (un)intended effects of persuasive attempts.

In a return to the focus on motivation, this dissertation seeks to further explain the roles of individual differences, social contexts and messages in inciting reactance and attempts to restore freedom, or lack thereof, to better predict responses to real-world persuasion.

Exploring Moderation Effects of Psychological Traits on the Experience of Reactance in the Context of COVID-19 Vaccine Mandates

Both clinical and persuasion psychology have been developing work on psychological reactance theory (PRT) for decades (Rosenberg & Siegel, 2018). A novel idea that emerged from clinical psychology has been to conceptualize reactance as an individual difference, referred to as trait reactance or reactance proneness (Brehm & Brehm, 1981; Rosenberg & Siegel, 2018). Communication research has borrowed this concept as an audience segmentation technique (Quick et al., 2013), but not without controversy. Due to validity issues identified with trait reactance, Rosenberg and Siegel (2018) called for the investigation of additional personality factors as predictors of reactance. An additional trait of interest in both the clinical and persuasion psychology domains of reactance research is locus of control (Levenson, 1981), which is a trait variable thought to reflect how much control of their circumstances any individual perceive they have. Locus of control has many of the same underpinnings as original theorizing about autonomy and learned helplessness from Brehm and Brehm (1981). Yet, locus of control has been understudied in the reactance literature writ large and has been particularly absent in studies related to mandates even though it has implications for activism (Levenson, 1981).

Therefore, this study seeks to incorporate locus of control and trait reactance as moderators of the mediation path between freedom threat and attitude through state reactance in a policy context. Specifically, the current study will test this model on messages related to COVID-19 vaccine mandates on U.S. college campuses.

Trait and State Reactance

The initial concept of trait reactance was not included in the original theory from Brehm (1966) but was subsequently adopted to the major theory as articulated by Brehm and Brehm (1981) after clinical psychology began exploring trait reactance (Rosenberg & Siegel, 2018). In contrast to the definition of state reactance as a mediating state between a threat and attempts to restore freedom (Brehm, 1966; Wicklund, 1974), trait reactance is conceptualized as a moderating variable driven by personality differences (Rosenberg & Siegel, 2018). Therefore, the logical implications in both clinical psychology and persuasion studies are to use trait reactance as an audience tailoring, or audience segmentation, method (Quick et al., 2013; Rosenberg & Siegel, 2018).

A series of measurement studies led to the development of the 11-item Hong Psychological Reactance Scale (HPRS) (Quick et al., 2013). Although there have been a multitude of attempts to measure trait reactance, Hong and Faedda's (1996) scale remains the most popular due to "the best conceptual correspondence with the reactance construct and the most favorable psychometric properties," (Dillard & Shen, 2005, p. 149). In response to greater consensus on the measurement of trait reactance, Quick et al. (2013) called for trait reactance to be considered as a main predictor and a moderating variable in persuasion research.

However, concerns about the validity of trait reactance and what role the variable should play have been raised consistently (Reynolds-Tylus, 2019; Rosenberg & Siegel, 2018). The core idea behind trait reactance is that "individuals differ in their needs for autonomy and self-determination," (Reynolds-Tylus, 2019; Wicklund, 1974). However, debates about the ability to measure reactance as a personality trait have hampered progress on understanding trait reactance (Quick et al., 2013; Rosenberg & Siegel, 2018). Further, most reactance studies measure trait

reactance or state reactance, instead of using both as suggested by Dillard and Shen (2005). Thus, the field has been unable to disentangle if trait reactance and state reactance are both useful predictors of key outcomes, including attempts to restore freedom by derogating messages, bolstering attitudes in the opposite direction of the advocated position, and engaging in discouraged or prohibited behaviors directly.

Some scholars have argued that there is no trait-level variable of reactance at all (Shoham et al., 2004; Silvia, 2006; Rosenberg & Siegel, 2018). Specifically, Silvia (2006) found that though individuals higher in trait reactance argued more for lower threat messages, they also tended to agree more with higher threat messages. However, the study was of less than 100 participants and the message manipulation was ill-defined.

Despite controversy, trait reactance has been found to be a useful predictor of engaging in risky health behaviors (Miller & Quick, 2010), as well as responses to persuasive messages and graphic health warning labels (Reynolds-Tylus, 2019). Furthermore, trait reactance has been shown to interact with message features to influence reactance (Dillard & Shen, 2005; Reynolds-Tylus, 2019).

Within the context of COVID-19 mandates, Albarracín and colleagues (2021) examined trait reactance as a moderator of the difference between free choice and vaccine mandates on individual's intentions to vaccinate and found no effect of trait reactance. However, over half the participants in the Qualtrics panel had already received at least one COVID-19 vaccine dose. The study was examining intentions to comply with future vaccine mandates, possibly outside of COVID-19. However, the high prevalence of vaccinated in the study makes the results difficult to interpret, since a mandate for an action an individual has already engaged in is possibly not perceived as a freedom threat. The study of COVID-19 vaccine mandates and state reactance

conducted by Kriss and colleagues (2022), which did find evidence of reactance to vaccine mandate messages, excluded participants who were already vaccinated from analyses related to reactance.

In a study seeking to predict intentions to engage in individual health behaviors to prevent COVID-19 infection and transmission, Horner and colleagues (2023) found that trait reactance and perceived threat from COVID-19 interacted such that greater trait reactance negatively predicted intentions to engage in mitigation behaviors unless the perceived threat from COVID-19 was high. Similarly, Young and colleagues (2022) used data from a national survey and found that trait reactance negatively predicted mask wearing in adults aged 18-49. Due to the weight of the evidence in favor of trait reactance, the first hypothesis is advanced.

H1: Trait reactance will moderate the relationship between perceived freedom threat and state reactance.

Reactance, Freedom Restoration, and Locus of Control

Early reactance scholars devoted much theorizing to explaining when individuals would seek to restore freedom directly or indirectly and to understand the conditions for when they *would not* (Brehm & Brehm, 1981; Wortman & Brehm, 1975). Wortman and Brehm (1975) theorized about perceptions of freedom, freedom threat and reactance being worn down through a lack of control in a model that combined reactance theory (Brehm, 1966) and the learned helplessness model (Seligman, 1974; 1975). See figure 1. Specifically, they theorized that reactance occurs under narrow conditions:

"Subjects who are exposed to uncontrollable outcomes should experience reactance only if they expected to have some control over the outcomes, and then only to the extent that the outcomes are important," (Wortman & Brehm, 1975, p. 307).

Ultimately, they expected that "reactance will precede helplessness for individuals who originally expect control," (p. 308). They advocated for a method known as "helplessness training" to test these relationships. Helplessness training refers to experimental exposure to "inescapable aversive outcomes or noncontingent positive outcomes," (p. 296). The integrated model from Wortman and Brehm (1975) never took hold and instead theorizing about the role of expected control and reactance would become Brehm's (1979) motivational theory and lead to broader work on an individual difference variable—locus of control (Brehm & Brehm, 1981).

Figure 1. Integrated model of reactance and learned helplessness reprinted from Wortman and Brehm (1975)
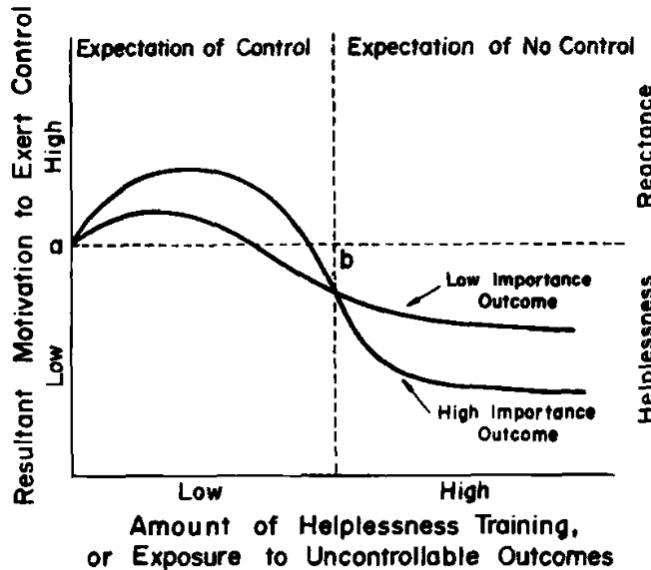
Fig. 1. The integrative model.

The variable locus of control is taken as synonymous with a personality variable that informs baseline perceptions of control (Brehm & Brehm, 1981). Originally, Rotter (1966) theorized locus of control as a continuum between internal and external locus of control such that those higher in internal control believe they have more control whereas those high in external control believe that that they are controlled by forces beyond themselves, including other people and chance. Theoretical advancement by Levenson (1981) moved locus of control from a unidimensional construct to a multidimensional construct and improved measurement of the construct. The construct of locus of control in the context of health are numerous, so much so that a specific scale was created for measuring locus of control in health contexts (Wallston, Wallston et al., 1978; Wallston, 2005). In the COVID-19 era, many studies examined locus of control as predictors of health behaviors (e.g. Cheng et al., 2016; Weinhardt & Ruckert, 2023; Origlio & Odar Stough, 2022). However, the research is much more limited in the context of reactance.

The first studies of locus of control and reactance come from Cherulnik and Citrin (1974) and Pittman and Pittman (1979). Both studies found evidence that individuals with greater internal locus of control experienced greater reactance in the face of restrictions, as expected. Concerningly, both studies were exceedingly small (approximately 100 students). In the wake of these studies, Brehm and Brehm (1981) incorporated the idea that internal locus of control, opposed to more external, react more strongly to influence attempts. More contemporary studies of the relationship between locus of control and reactance ought to be considered. In the clinical context, locus of control has been used to successfully predict medication adherence (De las Cuevas et al., 2014; De las Cuevas, 2023). Further, a recent review of the evidence found that

health locus of control and reactance interact to predict medication adherence (De las Cuevas, 2023).

In the context of persuasion, Xu (2017) showed that individuals higher in internal locus of control reacted more negatively to a controlling language message condition than did those with greater external locus of control. No other studies of locus control and reactance were identified. Locus of control may be an important moderator in the relationship between reactance and freedom restoration. Reasonably, as a greater sense of external locus of control means that individuals believe they have very little control over events (Rotter, 1966; Levenson, 1981), they may be less inclined to engage in direct (i.e. engage in prohibited behavior) or indirect (i.e. boomerang attitude change) freedom restoration.

Although locus of control studies run the gamut from applications in business, pro-environmentalism, medication adherence, and cancer survivorship, the construct may be particularly apt for studying reactance to messages in a policy mandate context, because it has been studied in the realm of political activism from its inception (Levenson, 1981). In a more recent study of the US and Germany, adults higher in internal locus of control were more likely to "contribute to climate change mitigation, to donate money and in-kind, gifts to charitable causes, to share money with others, to cast a vote in parliamentary elections, and to donate blood," (Andor et al., 2022). In addition, external locus of control has been historically thought to lead to less resistance and activism, though the results in the early empirical work were very mixed (Levenson, 1981). Considering the evidence that locus of control predicts responses to restrictions, the second hypothesis is advanced.

H2: Locus of control will moderate the relationship between reactance and attitude.

Reactance and Policy Mandates

Brehm and Brehm's (1981) defined freedom threats as *anything* that makes exercising a perceived freedom more difficult, which encompasses a wide array of phenomena. The two areas of psychology that have taken the most interest in reactance have been clinical psychology and health persuasion, which has subsequently driven the types of freedom threats investigated (Rosenberg & Siegel, 2018). Most reactance research has focused on overcoming resistance to psychiatric and psychological treatment, or preventing backlash to mediated persuasion attempts to change health behavior (Reynolds-Tylus, 2019; Rosenberg & Siegel, 2018). The fields that took up PRT may also explain a relative lack of research on actual restrictions and enforcements. However, the COVID-19 era and its aftermath have renewed interest in PRT as a framework for understanding (non)compliance and backlash at varying levels of persuasion and policy.

Some reactance studies during the COVID-19 era were focused on communication and persuasion outcomes, in addition to or instead of, policy. For example, McGuire and Ball (2022) used reactance to predict sharing of videos about COVID-19 guidelines and perceptions of COVID-19 as a public threat. Furthermore, some studies examined personal behavior intentions outside the mandate context. In the context of activist advertising, COVID-19 era pro-masking advertisements from brands *Uber* and *Heineken* were shown to incite reactance and reduce behavioral intentions whenever a tagline at the end of the ad was high threat, compared to the same advertisement without a freedom threatening tagline at the end (Shoenberger et al., 2021). Knapp and colleagues (2021) examined the effect of self- versus community-oriented messaging on reactance and social distancing behavior. Although they found no direct effect of the message manipulation, they did find that self-oriented messaging interacted with greater financial distress

to predict greater reactance, whereas those in the community-oriented message condition had the same amount of reactance, which was lower, across levels of financial distress. In a study of experimentally manipulated Facebook posts, Lu and Sun (2022) found that anti-vaccine comments and negative emoji reactions to pro-vaccine posts predicted reactance and that reactance predicted COVID-19 vaccine hesitancy. Interestingly, a follow-up experiment by Sun and Lu (2023) showed that rebuttals by the Centers for Disease Control and Prevention reduced reactance.

Another series of studies examined responses to behavioral mandates at the policy level. In the context of vaccination, Sprengholz and colleagues (2021) showed that reactance was generated by both mandates and scare supply. Specifically, mandates led to greater intentions to avoid COVID-19 and the chickenpox vaccine. In another series of studies of the US and Germany, Sprengholz et al. (2022) showed that vaccine mandate messages for COVID-19 incited reactance and that reactance to COVID-19 vaccine mandates was negatively associated with intentions to get a flu vaccine.

Across a series of studies featuring multiple COVID-19 era recommendations (e.g. staying home, standing far apart), Krpan and Dolan (2022) showed that commanding language decreased intentions to comply compared to non-commanding messages and that participants reported anger toward commanding messages. However, against expectations from PRT, commanding messages increased intentions to comply compared to control messages. Ball and Wozniak (2022) examined induced reactance by having survey participants write about a COVID-19 message they received, whether they agreed or not, and looked at several self-reported psychological predictors. They found that issue importance and message fatigue predicted freedom threat and subsequent reactance. Furthermore, they found that reactance was

negatively associated with adhering to social distancing guidelines and hygiene measures related to COVID-19. Ball and Wozniak (2022) also examined the differences between self-reported political parties on issue importance and message fatigue. Democrats reported higher issue importance and lower message fatigue than Independents and Republicans.

Studies on masking have focused largely on political ideology and reactance. Rains et al. (2022) showed that COVID-19 risk and political partisanship at the state level interacted to predict responses to the imposition and end of mask mandates. On a smaller scale, Dillard et al. (2023) similarly showed that anti-masking norms and political ideology predicted reactance in response to a pro-masking campaign on a university campus that was in a state with a mask mandate (Pennsylvania Department of Health, 2021). Further, they found that descriptive and injunctive norms each interacted with political ideology such that the influence of anti-masking norms on freedom threat was stronger for conservative students. Further, reactance, anti-making norms, and ideology all predicted masking. Considering the strong evidence for reactance in response to COVID-19 era messaging and policy, the third hypothesis is advanced.

H3: Reactance will mediate the relationship between freedom threat and attitude.

Individual Differences and Properties of Threat

Threats to freedom include anything that makes exercising a particular freedom more difficult (Quick et al., 2013). Reactance theory holds that the characteristics of freedom threats determine the amount of reactance and subsequent attempts at freedom restoration (Brehm, 1966; Brehm & Brehm, 1981). Two early dimensions were part of original theorizing: magnitude of threat and the difference between direct and indirect threats (Brehm & Brehm, 1981).

Threat Direction

Indirect threats have always been of interest to the theory, but the studies have been limited. Brehm (1966) hypothesized that "If the loss of a free behavior to an observed person could just as well happened to oneself, then one's own free behavior is threatened," (p. 7), and more contemporary scholars have coined this as 'vicarious reactance,' (Sittenthaler et al., 2016). The first empirical study of the phenomena of indirect threat and vicarious reactance came from Andreoli and colleagues (1974). More recently, Sittenthaler and colleagues (2016) showed that vicarious reactance does occur, but that it is distinct from reactance incited by threats directed at the self. They showed that vicarious reactance is processed more cognitively whereas threats to the self are processed more emotionally. Due to the theoretical tradition, and suggestive studies, reactance scholars have explicitly called for more study of indirect threats (Rosenberg & Siegel, 2018).

Magnitude of Threat

Brehm and Brehm (1981) clearly state that "the greater are the forces acting against exercising a freedom, the greater will be the perceived threat to that freedom" (p. 57). However, there has not been much testing of the predicted relationship between magnitude of threat and magnitude of reactance, perhaps largely because until recently most reactance research focused on clinical psychology or individual behavior change messages.

As this study is a secondary analysis of the message effects experiment from Kriss et al. (2022), those findings should be reviewed. In the original study, there was no difference in freedom threat perceptions between the indirect and direct threat conditions, but the higher

magnitude condition, which centered around sanctions, was perceived as more freedom threatening than the no sanctions condition. Further, an interaction emerged such that the most freedom threatening condition was the indirect threat with sanctions. As the original study used political ideology as a covariate, this secondary analysis thought the experimental conditions may interact with other identified individual difference variables: locus of control and reactance proneness. Thus, a series of research questions were tested.

RQ1: Is the relationship between direction of threat and magnitude of threat on perceived freedom threat moderated by external locus of control?

RQ2: Is the relationship between direction of threat and magnitude of threat on perceived freedom threat moderated by reactance proneness?

RQ3: Is the relationship between direction of threat and magnitude of threat attitude toward the mandate moderated by external locus of control?

RQ4: Is the relationship between direction of threat and magnitude of threat attitude toward the mandate moderated by reactance proneness?

Method

The present study is a secondary analysis of a message effects experiment conducted in 2021 which explores psychological traits as moderators of the reactance process. The main effects of the message experiment were reported by Kriss et al. (2022). Consistent with the principles of Open Science (Spellman et al., 2018), study materials are available publicly on the Open Science Foundation website:

https://osf.io/7dysq/?view_only=29935ccdeb7245199ffa62bcbe956279

Participants

Data were collected during Spring 2021 prior to the imposition of any mandates for COVID-19 vaccinations at study sites. Data was collected in two communication departments, one in a major university in the Southwestern U.S. and the other at a major university in the Midwestern U.S. The study obtained IRB approval from both universities.

Of the 554 participants who started the experiment, 41 were removed from the study for the following reasons: failing to indicate age 18 years or older, failing the attention check, or dropping out of the study before viewing an experimental message. The full sample of complete responses was $N = 513$. Participants who had already received one or both doses of a COVID-19 vaccine were removed from the sample, including those that failed to indicate vaccine status (n = 142). The final sample included in the analyses consisted of 371 participants. See Table 1 for a demographic description.

Design and Procedure

A 2 (threat directness: direct/indirect) x 2 (threat magnitude: sanctions present/sanctions not present) between-subjects control group design was employed. Participants first responded to some background questions. Next, participants were randomly assigned to a message. Finally, participants were asked to respond to a series of individual difference measures, reactance and persuasion outcome measures, and demographic measures. After all questions were answered, participants were debriefed on researcher deception.

Stimuli

The stimuli in this study were designed to look like official university communication regarding a potential COVID-19 vaccine mandate to take effect Sept. 1, 2021. To make the stimuli as externally valid as possible, the materials used the official visual identities for each university, including logos, fonts, colors, and letterheads. Specifically, the communication centered around starting the democratic process to amendment the student code of conduct at each school. The manipulation of direction of threat was achieved by having the policy be at the university of the participant or the university of the other site. Approximately half of participants read a message about a mandate at their own university, consistent with early research in reactance from Andreoli et al. (1974). The high magnitude of threat used a sanction, whereas the low magnitude of threat condition explicitly stated no sanctions would occur. Consistent with other university policies for other required vaccinations (Fawole et al., 2018), the sanction used was a registration hold in which students could lose their spot in courses. The control condition message was taken from CDC website and offered generic health advice regarding COVID-19.

Measures

Each item was measured on a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree), unless otherwise noted. See Table 2 for correlations among measured variables. Additional constructs were measured as part of the larger study; however, the methods section only includes measures used in the present analysis.

Reactance Proneness

An individual's degree of reactance proneness was assessed with the scale from Hong and Faedda (1996) which has 11 items. Respondents reported their agreement with a series of

statements (e.g. Advice and recommendations usually induce me to do just the opposite.), $M =$

3.62, $SD = 1.97$, $\alpha = .90$.

Table 1. Demographics of participants

Demographics	Number (%) of participants (N = 371)
Race/Ethnicity	
White	218 (58.8%)
Black or African American	49 (13.2%)
Hispanic	69 (18.6%)
Latin	16 (4.3%)
American Indian or Alaska Native	2 (.5%)
Asian	52 (14.0%)
Native Hawaiian or Pacific Islander (NHOPI)	4 (1.1%)
Other	8 (2.2%)
Age	
18-20	220 (59.3%)
21-24	135 (36.4%)
25-29	5 (1.3%)
30+	11 (3.0%)
Gender	
Man	135 (36.4%)
Woman	233 (62.8%)
Other/Prefer not to say	3 (.80%)
Political Party	
Democrat	177 (47.7%)
Republican	91 (24.5%)
Libertarian	24 (6.4%)
Green	9 (2.4%)
Other	67 (18.1%)

Note. Percentages may not add up to 100% because of missing data. The categories for race and ethnicity are not mutually exclusive.

External Locus of Control

The current study used a measure of external locus of control adapted from Levenson

(1981) that contained two items concerning the belief that powerful others control the world (e.g.

I feel like what happens in my life is mostly determined by powerful people.) and three items

that the world is governed by chance (e.g. It is not always wise for me to plan too far ahead

because many things turn out to be a matter of good or bad luck.). Although some researchers

use the locus of control scales (internal, powerful others, and chance) as independent variables,

belief in chance and powerful others are highly correlated and both reflect greater externality of

control (Levenson, 1981), therefore, the current study combined them within one scale, $M =$

3.11, $SD = .98$, $\alpha = .71$. For parsimony, the current study only used externality scores as

predictors of reactance, reflecting the idea from Rotter (1966) that internal and external exist on a

continuum.

Perceived Freedom Threat

The perception of freedom threat was assessed with four items (Dillard & Shen, 2005).

Respondents reported their agreement that the requirement tried to make a decision for them,

pressure them, threaten their freedom to choose, and manipulate them ($M = 4.00$, $SD = 1.94$, $\alpha =$

.93).

State Reactance

Reactance was measured using anger and negative cognitions (Dillard & Shen, 2005).

Participants indicated their level of anger in response to four items (i.e., irritated, aggravated,

annoyed, and angry) on a 1 to 7 scale. Negative cognitions were assessed with the following

stem, "The thoughts you had while reading the potential campus COVID-19 vaccine requirement

were mostly…" measured on a seven-point semantic differential scale (Quick et al., 2015) with

three bipolar adjective pairs (good/bad, favorable/unfavorable, and valuable/worthless). A

composite score of reactance was created by taking the average of both anger and negative cognition scores, $M = 3.62$, $SD = 1.97$, $\alpha = .98$.

Attitude Toward Campus Vaccine Mandate

Participants reported the degree to which their attitude (Fishbein & Ajzen, 2010) toward the potential campus COVID-19 vaccine requirement was bad, good, negative, and positive (M = 4.36, SD = 2.11, $\alpha = .98$).

Analysis Plan

Hypothesis Testing

Consistent with Psychological Reactance Theory, it was hypothesized that reactance would mediate the relationship between freedom threat and attitude toward the mandate. Additionally, reactance proneness, or trait reactance, would moderate the relationship between perceived freedom threat and experienced reactance (i.e. state reactance). Furthermore, we also hypothesized a moderation effect of a greater externalized locus of control on the relationship between reactance and attitude. See Figure 2 for a hypothesized model.

The current study tested the model (Hayes' Model 21) using the PROCESS macro in SPSS (Hayes, 2013). Following recommendations of Preacher and Hayes (2004), a non-parametric percentile bootstrapping (5000) procedure was utilized to compute a confidence interval around the indirect effect.

Research Questions

A series of general linear models were estimated to test interactions between the experimental conditions of direction of threat and magnitude of threat and individual difference

variables on reactance constructs for exploratory purposes. These analyses excluded the group

that saw a control message, thus the sample size for these tests was ($n = 297$). The predictor

variables were mean centered and included all individual predictors, two-way interactions and

the three-way interaction. There was a model which regressed perceived freedom threat on the

experimental conditions and external locus of control (RQ1), another model which regressed

perceived freedom threat on the experimental conditions and reactance proneness (RQ2).

Additionally, models were estimated that regressed attitude toward the mandate on the

experimental conditions and external locus of control (RQ3) and another model which regressed

attitude toward the mandate on the experimental conditions and reactance proneness (RQ4).

Figure 2

Hypothesized Model
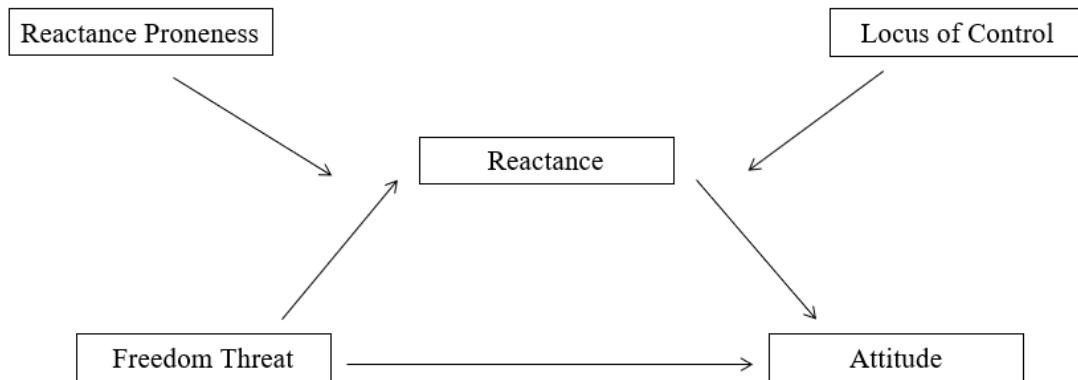
Note. The hypothesized model corresponds to Hayes (2013) Model 21.

Results

Hypothesis Testing

As expected, in support of H3, freedom threat predicted state reactance, b = .72, SE = .04,

$t(3, 367) = 19.22$, CI (0.65, 0.79), $p = .0007$. Also as expected, state reactance negatively

predicted attitude toward the vaccine mandate (H2), b = -.76, SE = .04, $t(4, 366)$, = -19.73, 95% CI (-.84, -.69), p < .0001. In addition, as expected, the indirect effect of freedom threat on attitudes toward the mandate was statistically significant in the model (H3), b = -55, SE = -.06, 95% CI (-.67, -.43). Against expectation, the direct effect of freedom threat on attitude toward the mandate was significant, b = -.23, SE = .04, $t(4, 366)$ = -5.89, 95% CI (-.31, -.15), p < .0001. The data supports a partial mediation model in which reactance mediates the relationship between freedom threat and attitudes toward vaccine mandate policies on campus.

Reactance proneness did positively predict reactance in a model controlling for freedom threat, b = .22, SE = .07, $t(3, 367)$ = 3.41, 95% CI = (.10, .35), p = .0007. Against expectation (H1), reactance proneness and freedom threat did not interact when predicting state reactance, b = .04, SE = .03, $t(3, 367)$ = 1.24, 95% CI = -.02, .10, p = .22.

Greater external locus of control did predict attitude in a model controlling for freedom threat and reactance such that those with more external locus of control had more favorable attitudes, b = .14, SE = .05, $t(4, 366)$ = 2.85, 95% CI (.04, .24), p = .0046. However, locus of control and state reactance did not interact to predict attitude toward the mandate, b = .01, SE = .03, $t(4, 366)$ = 0.42, 95% CI (-.04, .07), p = .67, against the expectation of H2.

The overall moderated-moderated mediation model showed no moderation effects of reactance proneness on the relationship between freedom threat and state reactance nor a moderation effect of external locus of control on the relationship between state reactance and attitude b = .004, SE = .001, 95% CI (-.002, .004). No significant interaction effects were observed. See figure 3 for a visual summary of hypotheses tested.

Research Questions

A series of exploratory analyses showed a variety of interactive patterns between individual differences and experimental conditions on perceived freedom threat and attitude toward the mandate. Locus of control did not moderate the relationship between the experimental conditions and perceived freedom threat (RQ1), $b = .44$, $t(289) = .77$, $p = .44$, *partial η^2* = .002. Locus of control did not predict perceived freedom threat, $b = .27$, $t(289) = 1.83$, $p = .07$, *partial η^2* = .01. Only the threat magnitude condition had a main effect on freedom threat such that those in the sanctions condition experienced greater freedom threat perceived than those exposed to the message without sanctions. $b = .67$, $t(289) = 3.06$, $p = .002$, *partial η^2* = .03. The interaction between conditions resulted in the no sanctions condition generating lower perceived freedom threat for an indirect rather than a direct threat $b = -1.04$, $t(289) = -2.83$, $p = .02$, *partial η^2* = .02. No other predictors were significant. See Figure 4.

In a model including interactions with conditions, reactance proneness independently predicted perceived freedom threat, such that those higher in reactance proneness perceived greater freedom threat, $b = .66$, $t(289) = 7.04$, $p < .0001$, *partial η^2* = .15. The interaction between conditions in a model with reactance proneness resulted in the no sanctions condition generating lower perceived freedom threat for an indirect rather than a direct threat $b = -.87$, $t(289) = -2.16$, $p = .03$, *partial η^2* = .02. No other predictors were significant. Further, reactance proneness did moderate the relationship between the conditions (RQ2), $b = -.75$, $t(289) = -1.99$, $p = .047$, *partial η^2* = .01. When the threat was direct, the relationship between reactance proneness and perceived freedom threat was stronger in the no sanctions condition compared to the sanctions condition. But, when the threat was indirect, the relationship between reactance

proneness and perceived freedom threat was weaker in the no sanctions condition compared to the sanctions condition. See Figure 5.

In a model including interactions with conditions, external locus of control independently predicted attitude toward the freedom threat such that those with greater external locus of control had a more negative attitude toward the mandate, $b = -.37$, $t(289) = -2.29$, $p = .02$, *partial $\eta^2 = $* .02. Further, locus of control interacted with the threat magnitude condition to predict attitude toward the mandate such that the negative relationship between external locus of control and attitude toward mandate is significantly weaker when there are no sanctions, $b = -.79$, $t(289) = -$ 2.41, $p = .02$, *partial η^2* = .02. The interaction between locus of control and both of the experimental conditions was not statistically significant (RQ3), $b = .62$, $t(289) = .96$, $p = .34$, *partial η^2* = .003. However, the trends in the data show that a greater external locus of control has a positive relationship with attitude when there are no sanctions and the threat is indirect, whereas the relationship between locus of control and attitude was negative when the mandate had sanctions and when the threat was direct. The experimental conditions failed to independently predict attitude toward the mandate in a model including locus of control and interaction terms. The interaction of the conditions was only marginally significant $b = .88$, $t(289) = 1.79$, $p = .07$, *partial η^2* = .01. See Figure 6.

In a model including interactions with conditions, reactance proneness independently predicted attitude toward the freedom threat such that those higher in reactance proneness perceived greater threat $b = -.67$, $t(289) = -6.13$, $p = .0001$, *partial η^2* = .12. Models with reactance proneness and experimental conditions predicting attitude showed that the experimental conditions failed to independently predict attitude toward the mandate and the interaction of the conditions predicting attitude was not significant, $b = .78$, $t(289) = 1.65$, $p = $

.099, *partial η^2* = .009. Reactance proneness did not moderate the relationship between

experimental conditions and attitude toward the mandate (RQ4), *b* = .55, *t*(289) = 1.24, *p* = .22,

partial η^2 = .

Table 2

Correlations Among Measured Variables (N = 371)

	M	SD	α	1	2	3	4	5
1. Freedom threat	4.00	1.94	.93	—				
2. State reactance	3.62	1.97	.98	.76***	—			
3. Attitude	4.36	2.11	.98	-.76***	-.88***	—		
4. Reactance proneness	3.62	1.10	.90	.43***	.43***	-.38***	—	
5. External locus of control	3.11	.98	.71	.01	-.04	.09	0.16**	—

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

Figure 3

Hypothesized Model with Results
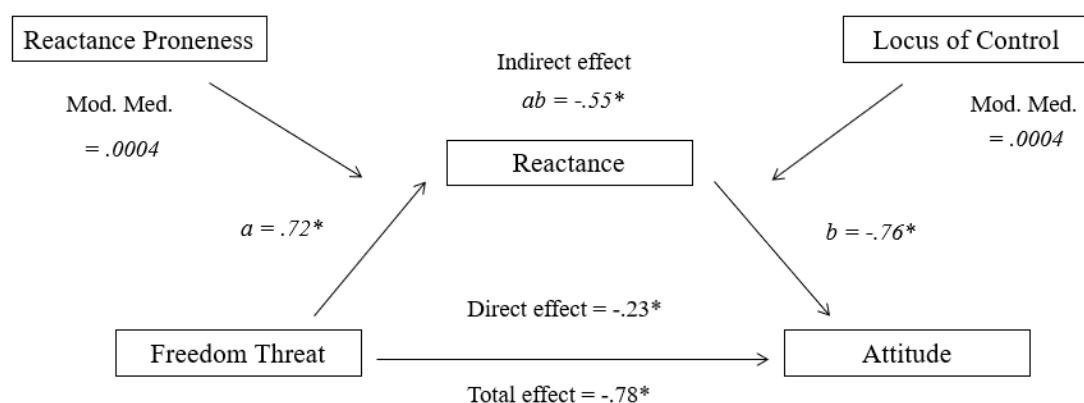
Note. *p < .001.

Figure 4

Interaction Among External Locus of Control, Direction of Threat and Magnitude of Threat on

Perceived Freedom Threat (RQ1)
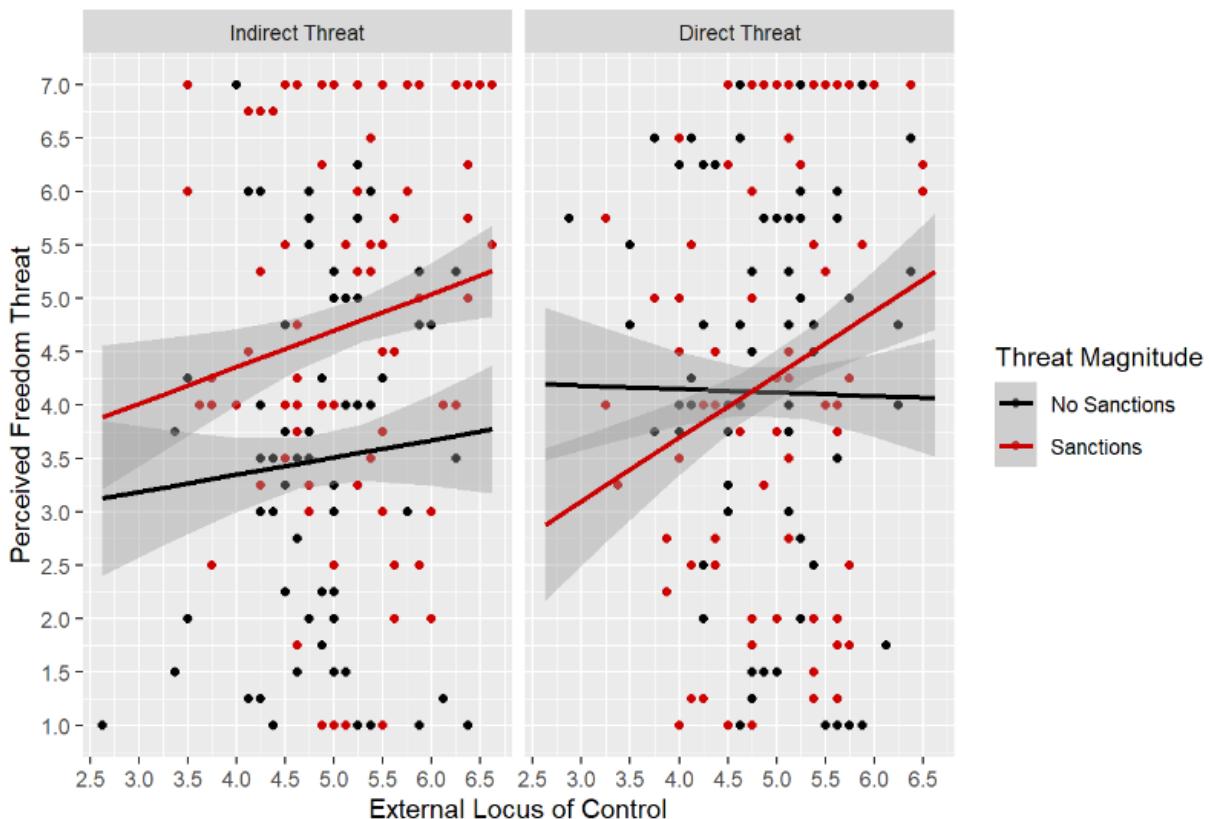
Note. Each dot represents an individual participant. Lines represented predictions of the model.

The shaded section represents the confidence interval of the prediction. Locus of control did not

moderate the relationship between the experimental conditions and perceived freedom threat.

Figure 5

Interaction Among Reactance Proneness, Direction of Threat and Magnitude of Threat on

Perceived Freedom Threat (RQ2)
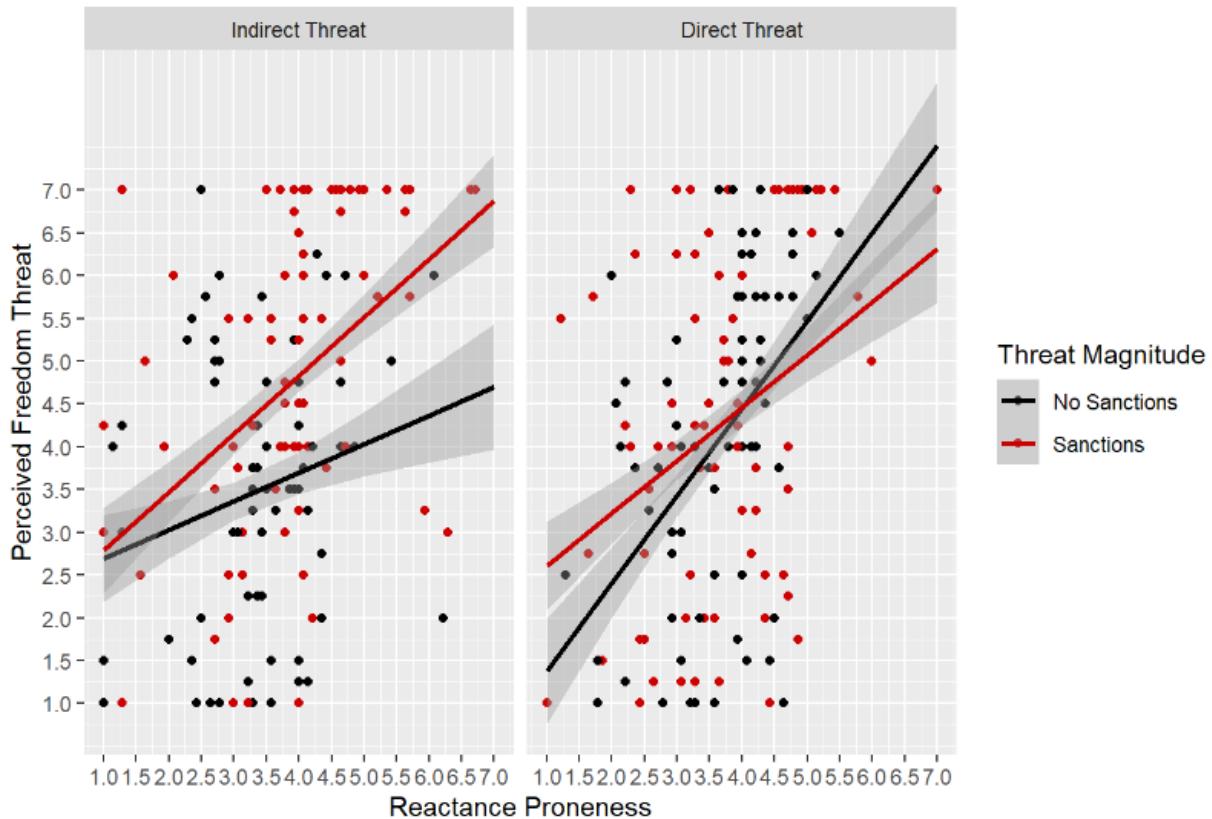
Note. Each dot represents an individual participant. Lines represented predictions of the model.

The shaded section represents the confidence interval of the prediction. Reactance proneness did

moderate the relationship between the conditions. When the threat was direct, the relationship

between reactance proneness and perceived freedom threat was stronger in the no sanctions

condition compared to the sanctions condition. In the indirect threat condition, the relationship

between reactance proneness and perceived freedom threat was weaker in the no sanctions

condition compared to the sanctions condition.

Figure 6

Interaction Among External Locus of Control, Direction of Threat and Magnitude of Threat on

Attitude Toward Mandate (RQ3)
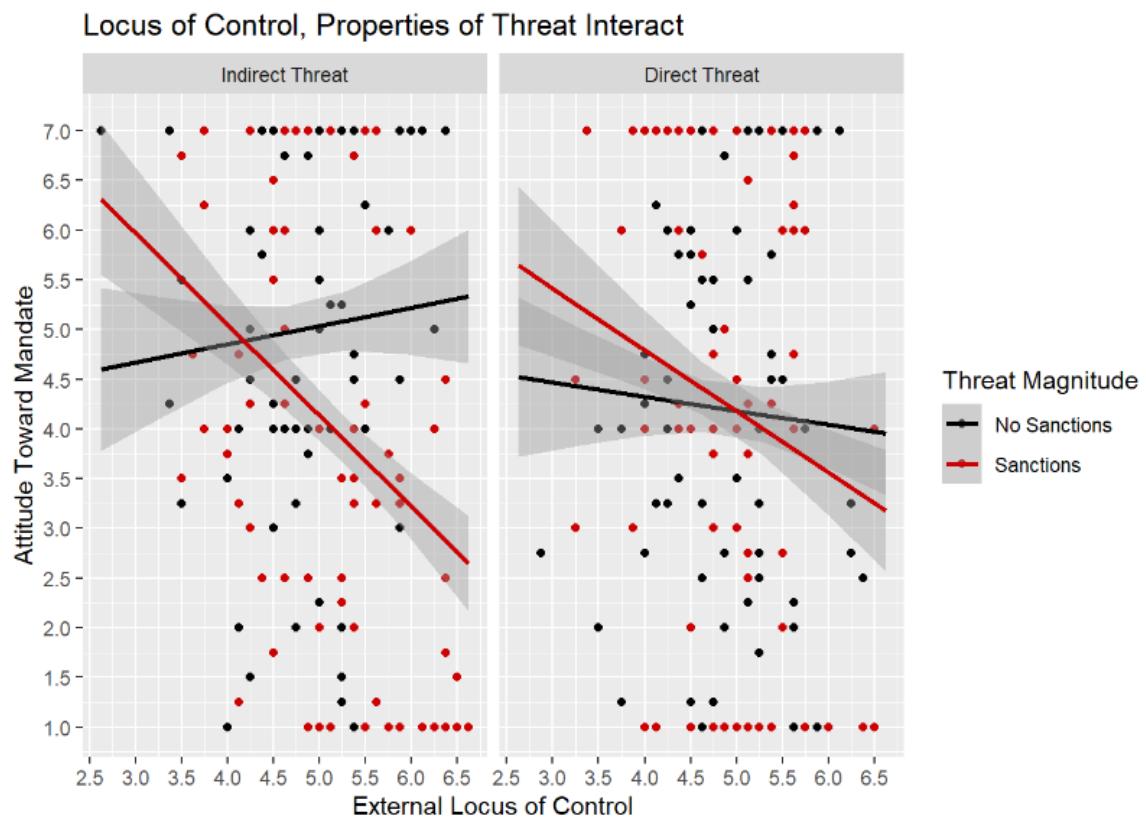
Note. Each dot represents an individual participant. Lines represented predictions of the model.

The shaded section represents the confidence interval of the prediction. Locus of Control did not

moderate the relationship between experimental conditions and attitude toward the mandate.

Figure 7

Interaction Among External Locus of Control, Direction of Threat and Magnitude of Threat on

Attitude Toward Mandate (RQ4)
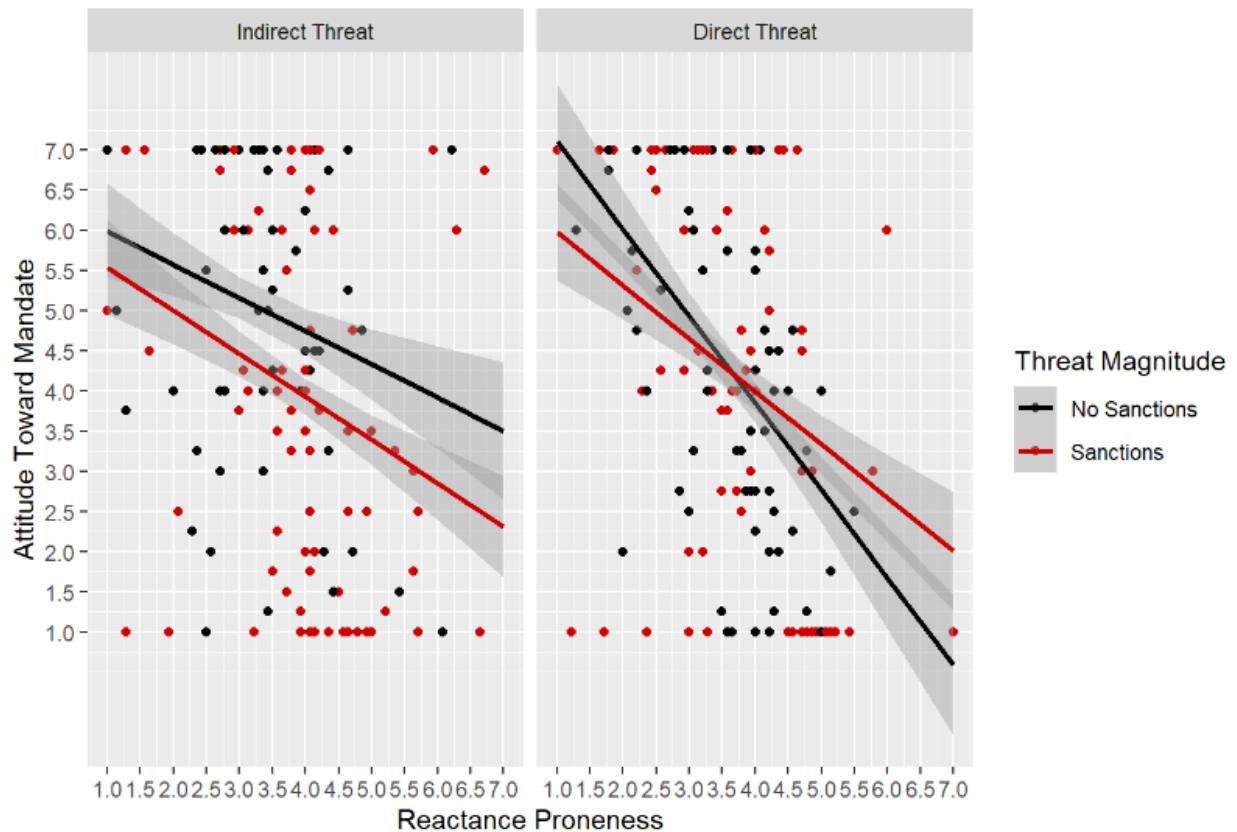
Note. Each dot represents an individual participant. Lines represented predictions of the model.

The shaded section represents the confidence interval of the prediction. Reactance proneness did

not moderate the relationship between experimental conditions and attitude toward the mandate.

Discussion

This study investigated the relationship between freedom threat and attitude toward a fictionalized COVID-19 vaccine mandate on two college campuses. It was hypothesized that greater perceived freedom threat would negatively predict attitude toward the mandate, mediated by reactance. Consistent with psychological reactance theory, we modeled that state reactance, a composite score of negative cognitions and anger, would mediate the relationship between perceived freedom threat and attitude toward the mandate (Rains, 2013). This study was primarily interested in psychological traits that might moderate the relationship between freedom threat, reactance and attitude. Specifically, this study explored whether reactance proneness, or trait reactance, would moderate the relationship between freedom threat and experienced reactance (i.e. state reactance) (Rosenberg & Siegel, 2018; Brehm & Brehm, 1981). Furthermore, the study also explored potential moderation effects of a greater externalized locus of control on the relationship between state reactance and attitude (Brehm & Brehm, 1981).

The study found evidence of a simple partial mediation in which freedom threat negatively predicts attitude toward a campus vaccine mandate independently, and also through an indirect effect on freedom threat from inciting reactance to negatively influence attitude. Two trait variables were explored as moderators of paths in the mediation. The moderated-moderated mediation model was not statistically significant. Ultimately, the data in this study is consistent with a simple mediation model only. In this study, the effect of freedom threat on attitude did not vary based on reactance proneness or external locus of control. The fact that neither personality variables moderated paths in the reactance process is inconsistent with original theorizing, but not necessarily inconsistent with the empirical evidence. The author knows of no studies that have tested the moderation paths specifically. Although the original writings would suggest that

freedom threats could interact with trait reactance to predict the magnitude of state reactance and that locus of control could interact with reactance to predict a freedom restoration behavior (Brehm & Brehm, 1981), the moderation of paths have not been tested previously. Early studies on reactance proneness and locus of control did not have the benefits of recent advancements in statistical analysis. For reactance proneness, most contemporary persuasion studies tend to use reactance proneness or state reactance as a predictor in persuasion studies. Dillard and Shen (2005), who advocated for the utility of trait and state reactance together, modelled trait reactance as a predictor of reactance and did not look at interactions with freedom threat. The studies of locus of control in reactance are extremely limited. The only studies identified looked at interactions between reactance and locus of control to predict medication adherence (De las Cuevas, 2023) and one study that modelled locus of control as interacting with message features to predict reactance (Xu, 2027). However, both constructs emerged as separate predictors in the model.

Reactance proneness predicted reactance when holding freedom threat constant. This can be interpreted as limited evidence that reactance proneness is a related, but independent, predictor of reactance. In the context of COVID-19 studies, most studies either use state (Ball & Wozniak, 2022; Dillard et al., 2023) or trait reactance (e.g. Albarracín et al., 2021; Horner et al., 2023; Young et al., 2022) as predictors. However, findings from this study may suggest that incorporating both constructs could reduce bias in predictions. Further, this would be a return to early theorizing by Dillard and Shen (2005).

Additionally, locus of control predicted attitude toward the vaccine mandate when holding freedom threat and state reactance constant such that those higher in external locus of control were predicted to have more favorable attitudes toward the vaccine mandate, consistent

with early theorizing on locus of control (e.g. Levenson, 1981). This is consistent with previous, albeit severely limited, research on locus of control and reactance for persuasion (Xu, 2017). This can be interpreted as limited evidence that future studies should consider including locus of control in models to reduce bias in the estimate of the effects of reactance on attitude change.

This study also investigated the interaction between individual differences, specifically a) external locus of control, b) reactance proneness and properties of freedom threats, c) threat direction, and d) threat magnitude to predict perceived freedom threat and attitude toward a campus vaccine mandate. Across four research questions which investigated interactions between the main experimental conditions reported by Kriss et al. (2020) and new individual difference variables, an inconsistent picture of findings emerged. Only reactance proneness moderated the interactive effect of the two property of threat conditions, direction and magnitude, to predict perceived freedom threat. When the threat was direct, the relationship between reactance proneness and perceived freedom threat was stronger in the no sanctions condition compared to the sanctions condition. In the indirect threat condition, the relationship between reactance proneness and perceived freedom threat was stronger in the sanctions condition compared to the no sanctions condition.

However, other main effects and two-way interactions emerged to predict perceived freedom threat and attitude toward the mandate. Most interestingly, locus of control was a negative predictor of attitude toward the mandate in models in which it interacted with the experimental conditions, which is opposite of the findings suggested by the main hypothesized model. And, though the three-way interaction was not statistically significant, the trends in the data show that a greater external locus of control had a positive relationship with attitude when

there were no sanctions and the threat was indirect, but the relationship between locus of control and attitude was negative when the mandate had sanctions and when the threat was direct.

Multiple comparisons, analytic flexibility and tests that are not preselected all increase the probability of findings being spurious (Ioannidis, 2005). Thus, all the results regarding individual differences and properties of threat interacting should be interpreted as hypothesis generating and grounds for future theorizing and testing, not as confirmatory findings.

Taken together, the results of this study suggest that reactance proneness and locus of control are trait variables that would improve estimates in reactance models and could be a useful audience segmentation strategy. Quick et al. (2013) called for investigating reactance proneness as a moderator, and Rosenberg and Siegel (2018) have called for investigating traits beyond reactance proneness. This study shows that both reactance proneness and locus of control predict responses to mandates. Thus, future studies should examine whether they are traits for tailoring messages by examining them as predictors of information seeking and avoidance, as suggested by Quick et al. (2013).

The results of this study, and past studies, suggest that individual differences are important predictors of reactance, particularly reactance proneness and locus of control. Thus, both can be thought of as ways of segmenting audiences to tailor strategies for persuasion. Recently, Richards et al. (2021) showed that the effectiveness of inoculation messages, freedom restoration postscripts and freedom threatening language were moderated by the level of reactance proneness. They offer a method for investigating the interaction between individuals and messages that can be replicated.

Limitations and Future Directions

The findings of this study must be considered in light of its limitations. From a methodological standpoint, the choice to use students as research subjects means that the sample is unrepresentative of the broader U.S. or world (Henrich et al., 2010). The relatively homogenous sample may explain the lack of stronger findings for the effect of external locus of control, as the variance across the sample was very low.

Further, using students recruited from the department of study which the research is about is an additional threat to validity because they are better at detecting hypotheses which can drive effects (Meltzer et al., 2012). To minimize this risk, this study measured a variety of constructs and include the use of deception, as suggested by Meltzer et al. (2012). In agreement with recommendations from Meltzer et al. (2012), we believe that student samples are not appropriate for understanding how phenomena would unfold in broader populations but that they are adequate for probing theoretical relationships. Furthermore, the study was designed to be relevant to the context of university students in the U.S.

The measure of locus of control in this study was abbreviated from Levenson (1981), which negatively impacted the reliability of the measure thus biasing estimates of effects. Future studies should attempt to use the full measure.

The analysis plan chosen also has limitations. The inferences that can be made from a mediation model using cross-sectional data are problematic due to an inability to assure the time-order of the hypothesized sequence. Thus, it is possible "the magnitude of the indirect effect via the mediator is substantially overestimated because the mediator and the outcome share omitted common causes," (Judd et al., 2014, p. 657). Even though the mediating role of reactance in the

relationship between freedom threat and reactance is well established in the PRT literature (Rains, 2013), future studies which measure these constructs at different time points would be valuable areas for further inquiry.

Future research should continue to explore the roles of reactance proneness and locus of control as predictors and moderators in the context of psychological reactance theory. In particular, Rosenberg and Siegel (2018) suggested that a useful alternative to trait reactance could be exploring other personality traits that could increase propensity for experiencing state reactance. Additionally, research in clinical psychology has probed interactions between reactance and locus of control (De las Cuevas, 2023), which suggests interactions between the two are worth exploring. Further, Xu (2017) tested locus of control as a factor that interacted with a message feature, which could be an avenue of further exploration for both locus of control and state reactance.

Conclusion

This study tested a moderated-moderated mediation model to identify individual differences that influence reactance. Trait reactance did not moderate the relationship between freedom threat and reactance but did independently predict reactance. Likewise, locus of control did not moderate the relationship between reactance and attitude, but independently predicted attitude toward a vaccine mandate. Overall, this study identified two useful covariates or audience segmentation variables that future studies of mandates can consider.

Studying Reactance in Social Settings: Coding Focus Groups for Valence and Expressed Emotion in Response to Public Service Announcements

Introduction

Television is a social medium. In 2024, 77% of Americans tuned in to the Super Bowl, the largest audience ever (Nielsen, 2024). The National Retail Federation estimated that more than half of 2024 Super Bowl viewers were throwing or attending watch parties, alongside a significant contingent who planned on going to a bar or restaurant to watch (National Retail Federation, 2024). The phenomenon of watch parties ranges from big national and global sporting events (Kim et al., 2020) to planned watch parties around entertainment TV programming such as *The Bachelor* (Nielsen, 2019). Additionally, new technologies, such as "WatchParty," "Scener" and group watch party features built into major streaming services, offer users the ability to have virtual watch parties with a small group of friends without time or physical proximity constraints (Ashworth, 2020).

Further, the rise of "Social TV" means that millions of solo viewers and group watch parties are interacting with others online about programming and advertising content (Nielsen, 2019). In a recent survey of Super Bowl watch parties, Kim et al. (2020) found that being at a watch party predicted engagement about the same program on social media.

Social viewing events drive major audiences and subsequent advertising revenue, including on prosocial issues. As many as 1 in 5 Americans watch the Super Bowl for the advertisements (National Retail Federation, 2024). Further, audiences have watched major controversies over programming and advertising go 'viral' on social media and news media. Yet

very few studies of message features for persuasion ever consider the differences between individual and group settings for media exposure.

Co-Viewing, Group Effects, and Psychological Reactance

There are three main reasons that psychological reactance scholarship has never examined group effects. First, the original theorists of PRT did not discuss the possibilities around reactance outside individual psychological and behavior change. Brehm and Brehm's (1981) theoretical explication around social power and reactance was limited. They only advanced theories about how individuals may respond to attempts to influence by individuals at varying levels of social power, and how interpersonal relationships (e.g. dyads) influence one another and generate reactance within the dyad. Second, the original PRT scholarship did not focus on mass-mediated influence. It wasn't until wave three of reactance scholarship, more than 20 years after reactance theory was formulated, that the field of mass communication adopted PRT as a framework of thinking about message features (Rosenberg & Siegel, 2018). Furthermore, because most communication research in PRT was applied to health communication and health behavior change (Quick et al., 2013), the focus on individual psychology remained. The final reason that social effects of reactance have been left unexplored is due to methodology. PRT was born out of experimental psychology. Experiments overwhelmingly focus on individual dynamics because they are easy to manipulate and control when designing experiments. However, PRT scholars have called for more research into the ability of intergroup contact to heighten or diminish reactance (Rosenberg & Siegel, 2018).

Therefore, the methods and rationale for this study draw from other areas of media effects and communication research alongside PRT to examine social effects of reactance in the context of mass-mediated persuasion.

Interpersonal Communication, Norms, and Psychological Reactance

The reasoned action approach posits that a variety of 'background factors' such as demographic variables, personality traits, and media exposure drive the formation of beliefs. In turn, these beliefs form the basis of a) attitudes, b) subjective norms, and c) perceived behavioral control, which then predict behavioral intention and behavior (Fishbein & Ajzen, 2010). See Figure 1 for an overview of the reasoned action model.

Interpersonal conversations may inform normative beliefs in ways that are predictive of behavior (Chung & Rimal, 2016; Rhodes et al., 2020). The literature on norms has focused on predicting behavior from norms at the individual level, but there is a need to understand how social dynamics influence the emergence, maintenance and evolution of norms (Legros & Cislaghi, 2020) which would be aided by qualitative investigation (Chung & Rimal, 2016)

Figure 1. The reasoned action model (Fishbein & Ajzen, 2010)
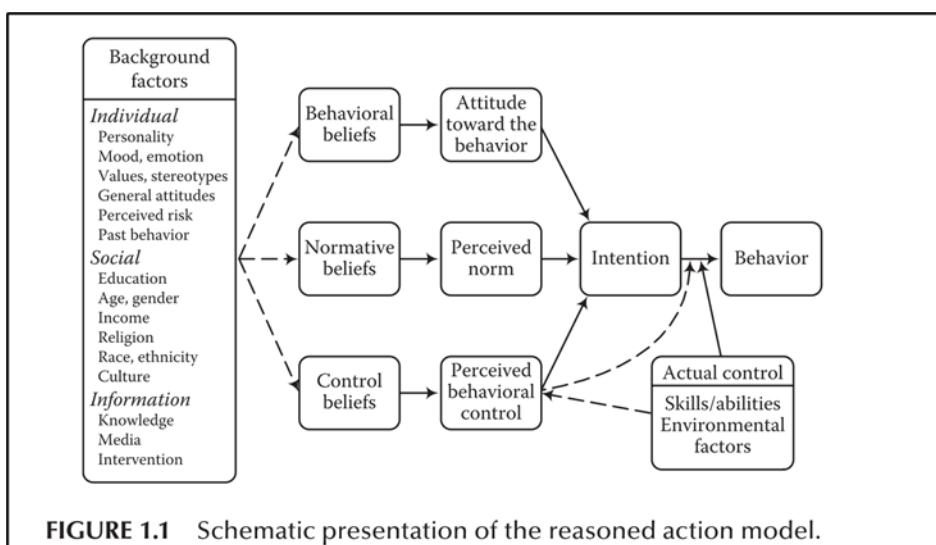
FIGURE 1.1   Schematic presentation of the reasoned action model.

Interpersonal communication, including group discussion, moderates the relationship between descriptive norms and behavior (Chung & Rimal, 2016). In a study about alcohol

consumption on campus, Real and Rimal (2007) found that students who reported talking to their friends about alcohol more frequently consumed more alcohol. Further, frequency of discussing alcohol interacted with descriptive norms such that the difference between perceiving lesser or greater frequency of drinking on campus was more predictive of consumption among students who talked to peers more frequently. Often, health campaigns even make interpersonal communication around a topic an explicit campaign goal (Southwell & Yzer, 2007).

In an over-arching review, Southwell and Yzer (2007) identified key roles that interpersonal communication may play in campaigns, a) unintended talk about messages, b) interpersonal communication as a mediator, or c) moderator of campaign effects, and finally d) as a planned outcome of campaigns. The function of interpersonal communication as an effective driver of persuasive effects has meta-analytic support. Jeong and Bae (2016) found that interpersonal conversation about campaigns, with and without campaign-generated conversation prompts, had 1.28 greater odds of influencing health behavior outcomes, stating "a look at the effect sizes shows that conversations are almost just as effective in bringing about behavioral outcomes as knowledge/awareness outcomes, despite the fact that behavior is the most difficult outcome to achieve," (p. 999). Although Jeong and Bae (2016) were looking at discussion as a moderator, Hornik and Yanovitzky (2003) proposed that interpersonal communication may mediate outcomes by changing normative perceptions (for the better, in the ideal case) or by expanding campaign reach through social diffusion (Southwell & Yzer, 2007).

However, interpersonal communication can function in counterproductive ways (Southwell & Yzer, 2007; Southwell, 2013). In one experiment, individuals assigned to chat online after watching a cannabis PSA exhibited more positive attitudes toward marijuana and greater normative pressure to use cannabis (David et al., 2006). These findings connect to the

failed $1 billion National Youth Anti-Drug Media Campaign pursued between 1998 and 2004. The speculated reasons for the campaign failure included a straight-forward reactance boomerang effect, but also what Cho and Salmon (2007) would describe as *social norming.* Specifically, "antidrug advertising conveys an implicit meta-message that drug use is commonplace," (Hornik et al., 2008, p. 2235).

One of the reasons that interpersonal norms and conversation surrounding media are difficult to predict is that the actual collective norm (e.g. how many students *actually* consume cannabis or actually approve of cannabis) is usually perceived incorrectly by individuals (Chung & Rimal, 2016). Further, the perceived norm is a more accurate predictor of individual behavior (Chung & Rimal, 2016). Social norms theory still struggles to understand the direction of the relationship between perceived norms and collective norms (Chung & Rimal, 2016; Legros & Cislaghi, 2020). Studies which explicitly manipulate the collective norm of the group could help separate out what reactions in social interaction would change if the collective norm were known (Chung & Rimal, 2016).

Message Features and Reactance

Although the potential opportunities of interpersonal communication in response to strategic communication have been established, there are still questions about what types of conversations lead to planned outcomes, and through what mechanisms those changes occur (Southwell & Yzer, 2007; Francis et al., 2021). Southwell and Yzer (2009) theorized that a) message function, b) message format, c) message timing, d) campaign topic, and e) network density factors influence conversations.

A core focus of the health communication phase of reactance research has been to identify message features to incite or diminish reactance (Rosenberg & Siegel, 2018), but unfortunately this work has yet to expand beyond understanding effects for individuals consuming messages alone. Message features that have been shown in some studies to diminish reactance at the individual level include narrative, empathy, provision of choice, and references to the consequences of others, as opposed to the self (Reynolds-Tylus, 2019). Further, appeals to norms are often used in health communication. A meta-analysis of appeals to norms in health communication showed that although normative appeals were effective, the effects were small and there is significant heterogeneity (Rhodes et al., 2020), similar to other message features (O'Keefe & Hoeken, 2021). Appeals to norms may incite reactance. In the context of voter registration, Reynolds-Tylus & Schill (2023) showed that appealing to a positive norm around being a registered voter incited reactance among those not registered and reduced intentions to register.

However, the message effects literature in reactance, similar to the wider persuasion literature, is limited by studies that only manipulate one or two message features and studies of interaction effects of message features are limited (O'Keefe & Hoeken, 2021). Therefore, most reactance studies rely on simple verbal or written persuasive messages. Reactance theory is infrequently used to evaluate complex stimulus, such as public service announcements.

Thought-Listing and Focus Groups in PRT

There is a precedent within PRT to examine naturally generated cognitions systematically as a way of measuring reactance. Reactance is modeled as a latent factor consistent of self-reported anger and negative cognitions. This model was first produced by Dillard and Shen (2005) and was further supported by meta-analytic model comparisons conducted by Rains

(2013) using 20 studies with almost 5,000 participants. This measurement of reactance remains the standard in the field (Ratcliff, 2021; Reynolds-Tylus, 2019; Rosenberg & Siegel, 2018). Both anger and negative cognition can be measured through administering thought-listing tasks. The thoughts are subsequently coded for anger and negative cognitions, by trained coders or by participants themselves (Dillard & Shen, 2005; Reynolds-Tylus et al., 2020). Thought-listing performs similarly to self-reported scale measures on negative cognitions in the context of PRT (Reynolds-Tylus et al., 2020). The main limitation of thought-listing, as used in experiments, is that it administered to individuals viewing messages alone, without social presence. However, the coding scheme associated with anger and negative cognitions could be applied to research methods outside of experiments with individuals, including focus groups.

The prevalence of focus group methods in the PRT domain is extremely low. Recently, Huang et al. (2024) cited PRT as a post-hoc explanation for some of the themes that came up in focus group discussions of COVID-19 mandates. However, only a few qualitative projects have started with PRT as an *a priori* framework for theoretical explanation. The studies that exist focus on experiencing reactance from message features. Youn and Kim (2019) investigated reactions to native advertising in interviews and focus groups. More recently, Ball and colleagues (2023) used focus groups to develop autonomy-threatening messages, which were subsequently used in a traditional message effects experiment. Like most research using focus groups, neither of these projects attempted to analyze group or interaction effects but instead only looked at aggregated individual analysis (Cyr, 2016). Applying though-listing coding procedures from previous studies to focus groups would provide a novel, but theoretically grounded, approach to measuring reactance generated within groups.

Individual and Group Effects in Focus Groups

In focus group studies, there are three levels of analysis: the individual, the group, and the interaction, which offer different insights (Cyr, 2016). In most marketing and health communication research, the individual unit of analysis is prioritized over group level and interaction analysis (Cyr, 2016). This leaves insights into social processes under-explored (Cyr, 2016; Hollander, 2004). Specifically, Hollander (2004) advocated for considering group characteristics and discussion patterns more central to focus group analyses:

> "…multiple and overlapping contexts foster both problematic silences (lack of disclosure) and problematic speech (strategic shaping of comments) in group discussions. These processes limit the usefulness of focus groups as a tool for understanding individual thoughts, feelings, or experiences. However, they make focus groups an excellent site for analyzing the processes of social interaction," (p. 603).

Comparison and analysis at the group level is worth undertaking as media viewing can largely be a social activity, which can influence cognition. The author knows of no studies within the psychological reactance literature that examines group effects from focus groups. However, current projects in experimental literature have generated insights into social components of reactance. Kriss et al. (2022) found that both perceived freedom threat and reactance were highest among students who heard about a hypothetical vaccine mandate with sanctions at a college that was not their own. So, there is evidence an individual can feel threatened by perceiving others having a freedom threatened. Also recently, Rains et al. (2022) found relationships between state-level political ideology and disease risk on mask wearing during state-wide mask mandate imposition and removal. The study by Rains et al. was at a higher level than a group, but also speaks to need to consider social context in the study of PRT. Graupmann

et al. (2012) found evidence that being categorized as part of a group could incite reactance, and Rosenberg and Siegel (2018) speculated about the potential effect of the knowledge of an outgroup member influencing reactance.

RQ1: Will the proportion of a) positive, b) negative, and c) neutral thoughts vary based on whether group norms are known versus unknown?

RQ2: Will the proportion of emotions expressed across a) anger, and b) fear, vary based on whether group norms are known versus unknown?

Method

Consistent with the principles of open science (Spellman et al., 2018), all materials needed to reproduce this study are publicly available on the Open Science Foundation website: https://osf.io/6djuf/?view_only=5af32fa157f4433e8ff022f3126526cf

Data Collection Procedure

Students from a journalism and mass communication department at a major university in the American Midwest were recruited to participate in focus groups about public service announcements regarding using cannabis. Six in-person focus groups took place in Winter 2022 among students recruited from communication courses (N = 39). An additional three focus groups took place in Spring 2023 which specifically screened participants (N = 12) for current cannabis users using the CUDIT instrument (Adamson & Sellman, 2003). Spring 2023 participants were recruited for being cannabis users and were aware that the focus groups were made up of other cannabis users. In contrast, the focus groups in Fall 2022 were made up of a general student audience and participants only knew the topic. Therefore, in the groups that were screened for cannabis, the collective behavioral norm was known, whereas in unscreened groups

the participants had no information about the behavioral norm. See Table 1 for a description of recruitment materials.

All focus groups viewed the same four PSAs in a random order before having a broader discussion about driving under the influence of cannabis. In some groups, the broader discussion about driving under the influence of cannabis happened first. All participants were compensated with extra credit. All focus groups were approximately one hour and utilized the same moderator.

Table 1
Recruitment materials sent to journalism and mass communication students

	Abbreviated recruitment message
Fall 2022 (No exclusion criteria)	You are being asked to be a participant in a research study about attitudes toward health campaigns about driving under the influence. You have been asked to participate in this research because you are our priority audience for this campaign.
Spring 2022 (Cannabis users exclusively)	We are seeking participants for a series of virtual focus groups. These groups are intended to gauge young adults' responses to some health messages about cannabis use. To qualify for participation, you will be asked a series of questions about your cannabis consumption.

Note. Full recruitment materials are available at OSF.

Stimuli

The videos chosen for this study come from PSAs produced and aired in territories where cannabis is legal but driving under the influence of cannabis is a crime. See Table 2 for a description of the PSAs used in this study. The PSAs all come from established major campaigns. Two PSAs came from the drug impaired driving campaign created by Ad Council and the National Highway Traffic Safety Administration (NHTSA) (nhtsa.gov/risky-driving/drug-impaired-driving). The other two PSAs were from the Colorado Department of Transportation ongoing drugged driving campaign.

Table 2

Description of PSA Stimuli Used in Focus Groups

Public Service Announcement	Synopsis
"Basketball"	This ad was part of the "Drive High, get a DUI" campaign from the State of Colorado. A group of middle-aged men are playing a game of pick-up basketball. One is hyper-fixated on dribbling the ball and forgets to shoot the ball. The main copy is "Playing ball high is now legal. Driving to see the pros play afterward isn't."
"Party"	This ad was from NHTSA and Ad Council campaign. The ad shows an extended party scene then cut to a group driving home. The driver almost hits a woman, and she angrily hits the top of his car with her hands. It ends with "If you feel different, you drive different."
"Pineapple"	This ad was from NHTSA and Ad Council campaign. A couple is getting high at home. The man decides he's too high to cut a pineapple. He also decides he's too high to drive after the woman asks. The main copy is "Nope, I'm high. Let's order in." It ends with "If you feel different, you drive different."
"Roll one"	This ad was part of the "Drive High, get a DUI" campaign from the State of Colorado. The ad focuses on showing a car crash existing within a joint being rolled by two hands. The main copy is "Roll one now, roll one over later."

Note. All stimuli were 30-second-long video advertisements. Full stimuli available at OSF.

Moderator guide

The focus group questionnaire was designed to be semi-structured in accordance with guidelines from Tracy (2013). For each PSA, participants responded to probes regarding their initial experience of the ad, including feelings and thoughts and how the PSA could be approved.

All participants were asked the same questions about each PSA, which included "What feelings did you have while watching that PSA?" and "What thoughts did you have while watching that PSA?" They were also asked "Do you think this PSA would be effective with an audience like you?"

Codebook development and coding procedure

This study uses an iterative analysis approach as described by Tracy (2013), as opposed to grounded theory. The authors went through data immersion and primary cycle coding before breaking up the data into though units. All nine transcripts of focus groups were broken into thought units by the first author. The initial code sheet was devised by the first and second author. All thought units (N = 1093) were systematically scored by the first and second author to evaluate reliability.

Across nine focus groups of general student audiences and students screened for cannabis use, 51 participants generated $N = 1,093$ thoughts. The codebook captured the valence of thoughts, based on early reactance research done by Dillard and Shen (2005). Additionally, in accordance with Dillard and Shen (2005), the research team searched for affective category words based on the list created by (Shaver et al., 1987). The research team also looked for fear words using the list from Shaver et al. but added "scare" as a keyword for fear. The research team also chose to exclude 'dislike' as an indicator of anger due to the evaluative nature of the focus group environment. See Table 3 for more information.

Ultimately, $N = 1086$ thoughts were subjected to content analysis after thoughts that were clearly uninterpretable or ranked all four ads in one thought were removed. Both the first and second authors unanimously agreed on exclusions.

This study defined reliability as reproducibility of codebook. Reproducibility of the codebook is not the only definition of reliability used in content analysis in communication research (Hayes & Krippendorff, 2007). For example, Cohen's κ, a once popular measure of reliability in communication research, assess the statistical independence of the coders (Hayes &

Krippendorff, 2007). However, this study was investigating replicability of the codebook, not in the replicability of the coders chosen for the project (Krippendorff, 2004, 2016, 2019). Therefore, this study uses Krippendorff's α.

Further, even though many content analyses use multiple coders to assess reliability to then divvy up the remaining units of analysis excluded from reliability calculations among coders (e.g. Yang et al., 2023; Quick et al., 2023), the reason for splitting entire data sets between coders is merely practical. Since this study defined reliability as reproducibility of codebook with the appropriate statistical test, there was no need for all coders to assess all the units or for units to be divided evenly between all coders. Once reproducibility was established, the first author's judgments were retained for the entire data set.

Due to the low prevalence of emotion words associated with fear, and anger, both coders decided to code emotions based on unanimous agreement. Any statement that included an emotion keyword was included as expressing the emotion unless the comment was in reference to a) messages not included in the study (e.g. "*That's very true because I think about comparing this, the orange anti-nicotine campaigns and those are very visceral. They show really gross images and make people feel either scared or disgusted. And I feel like this, I mean obviously is really different because it's kind of trying a more comical appeal*," Thought 24). Statements with emotion keywords were also excluded if they explicitly stated they were not feeling an emotion mentioned (e.g. "*You don't really feel scared during it.*", Thought 248). Only coding for valence was subject to reliability testing.

Table 3

Description of the Codebook

Component	Description
Valence	
Negative	Thoughts that a) expressed disagreement with the message, b) negative thoughts toward the message, c) negative thoughts toward the message source, d) negative thoughts toward the advocacy, e) negative intention to comply with the advocacy, f) intention to engage in the risky behavior.
Neutral	Thoughts that expressed non-evaluative responses to the message.
Positive	Thoughts that a) expressed agreement with the message, b) self-identification, c) positive thoughts toward the message, d) positive thoughts toward the message source, e) intention to comply with the advocacy in the message
Emotion	
Anger	Keywords: aggravation, irritation, agitation, annoyance, grumpiness, exasperation, frustration, anger, rage, outrage, fury, wrath, hostility, ferocity, bitterness, hate, loathing, scorn, spite, vengefulness, resentment, disgust, revulsion, contempt, envy, jealousy, torment
Fear	Keywords: alarm, shock, fear, fright, horror, terror, panic, hysteria, mortification, anxiety, nervousness, tenseness, uneasiness, apprehension, worry, distress, dread

Data Analysis Procedure

Reproducibility of the Codebook

Krippendorff's (1970) α was utilized to determine if there was sufficient agreement between the two coders. All reliability measures attempt to determine the agreement between coders observing units (i.e. thoughts, clips etc.) independently of each other (Hayes & Krippendorff, 2007). However, Krippendorff's α is the preferred reliability measure for content analysis for practical and theoretical reasons. On the practical side, Krippendorff's α is a good standard measure to compare across content analyses because it can be used to analyze variables at any level of measurement, accommodate more than two coders, and retains validity in the presence of missing data and missing categories or scale points (e.g. categories originally created

in the codebook but never used by a coder) (Hayes & Krippendorff, 2007). The measure is also seen as methodologically superior because it corrects several shortcomings of other reliability measures. In addition to being practical, the ability to account for only the categories which show up in the data means that Krippendorff's α "is not biased by the difference between what the authors of the coding instructions imagined the data may be like, and what the data turned out to be," (Hayes & Krippendorff, 2007, pp. 79).

Unlike simple measures of reliability like percent agreement, Krippendorf's α has a clear interpretation in that a score of zero reflects a state in which "the units of analysis bear no statistical relation to how they end up being identified, coded, or described," (p. 79). Krippendorf's α is a sample statistic, meaning that the true α can only be calculated by using a bootstrapping procedure to estimate the sampling distribution. Therefore, this study used the KALPHA macro for SPSS with a bootstrapping procedure from Hayes and Krippendorff (2007) to calculate estimates of the α coefficient.

In accordance with guidelines from Krippendorff (2004), $\alpha > .80$ was interpreted as reliable (pp. 241-242). The reliability between the coders on the valence of thoughts was considered reliable, $\alpha = .89$, 95% CI [.87, .92]. If 1000 different samples were collected, the probability that different coders using the study codebook would score $\alpha = .90$ is 71%, but the probability they score below $\alpha = .80$ is less than 1%.

Descriptive Analysis of Differences Between Groups and Stimuli

As the data for this study is nested and non-independent, a descriptive analysis approach was used. Further, in accordance with guidelines for qualitative research (Tracy, 2013), there was a concerted effort to highlight quotes representative of themes identified in the focus groups.

Results

RQ1: Will the proportion of a) positive, b) negative, and c) neutral thoughts vary based on whether group norms are known versus unknown?

Groups that were screened for cannabis use had a different proportion of positive, negative, and neutral thoughts than groups that were not screened for cannabis use.

The unscreened groups had a greater proportion of negative thoughts (51.1%) than did the screened groups (32.9%). The unscreened groups had a smaller proportion of neutral thoughts (21.7%) than did the screened groups (39.5%). The proportion of positive thoughts was similar between unscreened groups (27.2%) and screened groups (27.6%). See Figure 2.

RQ2: Will the proportion of emotions expressed across a) anger, and b) fear, vary based on whether group norms are known versus unknown?

There was a very low proportion of thoughts that contained anger, fear, and surprise. Specifically, fear was found in only 1.6% of thoughts (n = 17), and anger was found in only 0.6% (n = 7). When participants expressed anger, it was mostly related to annoyance.

"Yeah. I feel like the second one [Car Crash Joint] was the closest I got to feeling lectured or being annoyed with it. Okay, stop preaching at me… so I think that's the last two ["Party" and "Car Crash Joint"] were more effective and not super preachy, if that makes sense," (Thought 311, Group 6).

Anecdotally, the most thoughts that contained anger generated by any group was 3 (group 5). Most of the thoughts with fear were generated by group 4 (41.2%, n = 7) and group 6 (35.29%, n = 6). Interestingly, no groups screened for cannabis use generated angry thoughts.

Also, only two of the seventeen thoughts that mentioned fear came from groups that were screened for cannabis.

Figure 2

Difference in Valences of Thoughts Between Screened and Unscreened Groups

	Unscreened	Screened
■ Positive	229	67
☐ Neutral	183	96
■ Negative	431	80
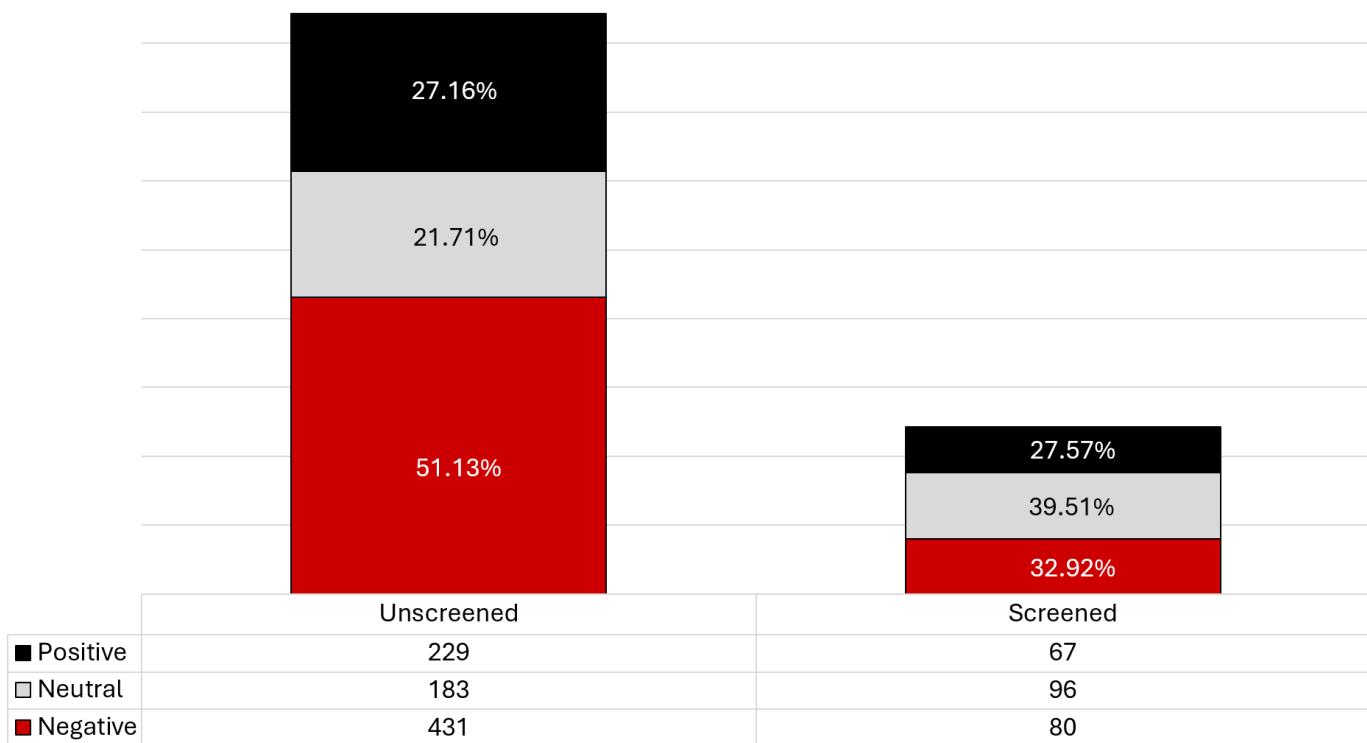
Note. Rounded percentages may not add up to 100%.

Exploratory Analysis

Considering the surprising lack of emotions found in the data, the research team also decided to explore the data for signs of perceived freedom threat. Although no reactance studies have coded thoughts for perceived freedom threat, the primary measure of perceived freedom threat has four Likert-type scale items: "The message threatened my freedom to choose," "The message tried to make a decision for me," "The message tried to manipulate me," and "The message tried to pressure me," (Dillard & Shen, 2005). Therefore, thought units were scanned for the words "threat," "freedom," "manipulate," "pressure" and derivatives (e.g. "free").

Zero thoughts mentioned freedom or manipulation. Four thoughts mentioned "threat." Two thoughts mentioned "pressure." All thoughts related to perceived freedom threat conveyed that the advertisements were not explaining the threat of driving under the influence of cannabis enough.

"One thing that I did think was maybe a little ineffective about this one is that…the other ones were more threatening, like you can get a DUI, you can crash and hurt people. This one was just more like, "Oh well yeah I probably shouldn't.'" (Thought 816, Group 6, Advertisement "Pineapple").

This was also true in the groups screened for cannabis use. The only other thought mentioning threat came from a group screened for cannabis use. The next two of thoughts, Thought 1009 and 1010 agreed, also from participants in groups screened for cannabis use.

"And obviously the goal of these advertisements are to get people to not drive high and get them to realize the threat of driving high. In the first advertisement, I don't think the reason I

gave it six as well is because I don't think it address the severity of the goal they're trying to accomplish," (Thought 1008, Group 2, Advertisement: "Car Crash Joint").

The same dynamic applied to the word pressure.

"Put it in a scenario where you can see the pressure and model that good behavior. Because, sitting at home and doing random small things, I feel like that's not as much of a scenario that could create that danger," (Thought 236, Group 2, Advertisement "Pineapple"). The other thought mentioning pressure expressed positive sentiments about advertising to dissuade the influence of *peer pressure*. "I also liked how there was a sense of somebody's not going to take 'peer pressure,'" (Thought 663, Group 5, Advertisement "Pineapple").

Discussion

This study pilot tested a novel approach to analyzing focus group data in the context of psychological reactance theory, and in all message effects and persuasion more broadly. This study used analytical techniques associated with individual thought-listing and applied them to focus group data. In contrast to much of the research within the field of reactance theory, this study also tested reactions to professional-produced mass market public service announcements. This study offers multiple insights both theoretical and applied.

First, the results indicated that groups did generate different patterns of cognition. Specifically, groups varied widely in the proportion of positive, negative, and neutral sentiments expressed. Some, but certainly not all, of the group variation could be explained by differences in group make-up. Specifically, groups that were screened for cannabis use, thus made up of cannabis users and made up of participants in which everyone was aware of the collective behavioral norm, expressed different proportions of sentiments than did the groups who were not screened for cannabis use, where the behavioral norm is unclear. Specifically, although both

screened and unscreened groups generated a similar proportion of positive thoughts, the groups screened for cannabis use generated a lower proportion of negative thoughts and a greater proportion of neutral thoughts than did the unscreened groups. Additionally, although the prevalence of emotions was extremely low, it is interesting that only the groups not screened for cannabis use generated any anger, and that they also generated almost all of the fearful thoughts.

One potential explanation for these differences is that, presumably, the PSAs were designed to appeal to cannabis users, thus groups that contain at least some non-users found these ads less personally relevant and less persuasive than groups made up exclusively of the cannabis users. This would be similar to other co-viewing studies (Tal-Or, 2016; Mora, 2016).

Another explanation for the difference could be that knowledge of the behavioral norm influenced the social dynamics within the groups. Group discussions are influenced by contextual factors related to group make-up (Hollander, 2004). Perhaps the group of cannabis users participating in groups they knew had other cannabis users exclusively led them to be more honest about their responses to the PSAs, whereas participants in the unscreened group may have been trying to intuit the cannabis use status of others and act in a way they perceived as socially desirable. Some reactance scholarship has shown that perceiving the presence of being in the ingroup reduces reactance (Graupmann et al., 2012), thus a group made up exclusively of cannabis users where everyone knew they were all cannabis users may have felt less reactance. This could be even more true as this study took place in a state where cannabis use is criminalized. A study which explores the differences between general audiences and audiences made up of cannabis users exclusively, which also systematically manipulates the knowledge participants have about the collective behavioral norm (i.e. cannabis use) of co-participants could disentangle this possibility.

However, the comparison of screened and unscreened groups may be confounded by the differences in group size. The groups that were screened for cannabis users were, by no intention of the research team, smaller than those groups that were unscreened. The screened groups were smaller and thus generated fewer thoughts. And, although comparing proportions corrected for the difference in the number of thoughts generated by each group, there could be a causal effect between group size and the overall sentiments of thoughts expressed. Future research which systematically manipulates group size and group make-up could help quantify whether focus group size is a confounder in the study of message effects.

The lack of expressed emotions toward the PSAs is remarkable. The lack of anger, a key marker of reactance, from the discussion of PSAs raises important concerns about how much reactance is generated from communication that has not been artificially manipulated to test reactance constructs. Although the PRT literature explicitly recognizes that many messages do not generate reactance due to either a low perceived value of the freedom or low perceived restriction from the threat (Brehm & Brehm, 1981; Kriss et al., 2022), it is curious that negative cognitions varied substantially, but anger did not. If reactance was experienced in this study, we might have expected greater co-variance. Further, the exploratory examination of signs for perceived freedom threat yielded the insight that when threat came up, it was in the context of advertisements being perceived as not threatening *enough*. The reactance field should embrace the calls for larger, multi-message studies of naturally occurring stimuli to determine the prevalence of reactance responses in the context of prosocial messaging (O'Keefe, 2015; Slater et al., 2015; Yang et al., 2024).

Limitations and Future Directions

The potential insights of this study should be considered in the context of the limitations imposed by the study design. The focus group protocol of the study guaranteed anonymity and did not require participants to disclose information about themselves. A future research effort could more systematically explore differences in group make-up and how it influences responses to PSAs by linking personal information, such as demographics and personality measures, to participants in focus groups in a method similar to Hollander (2004). Further, as previously mentioned, the results of this study are potentially confounded by the difference in group size. Future research should systematically examine group size in any replications of this procedure.

Focus groups are, an improvement in external validity because they capture social responses to messages. However, the focus group environment is still highly artificial. Future studies would benefit from methods that more closely approximate natural viewing, such as an in-home simulation lab. The generalizability of this study is limited due to the sampling method. Like most research on PRT, the sample is made up exclusively of college students (Reynolds-Tylus, 2019). Thus, the results should not used to predict how any other audience may respond to messages.

Conclusion

This study piloted a novel approach for measuring reactance in the context of co-viewing media. Results indicated that groups varied in the proportion of sentiments expressed toward public service announcements against driving under the influence of cannabis. Groups made up exclusively of cannabis users, which knew the behavioral norm of the group, generated fewer negative thoughts, no angry thoughts and fewer fearful thoughts compared to groups without knowledge of the behavioral norm. These findings suggest that the difference between individual

and co-viewing scenarios could moderate persuasive effects. Further, the low proportion of anger

and perceived freedom threat generated toward the public service announcements suggests that

externally valid stimuli may incite less reactance.

Do Warning Labels Mitigate Persuasion or Incite Reactance? An fMRI Experiment Among

At-Risk Young Adults

Reactance is a motivational state in which individuals seek to restore a threatened freedom. And the original theorists took the experience of reactance to be a "black box" that could never be measured (Brehm & Brehm, 1981; Dillard & Shen, 2005; Quick et al., 2013). Instead, they believed that the best evidence of reactance was measures of freedom-restoring cognitions and actions. Communication researchers challenged the original theory (Brehm, 1966; Brehm & Brehm, 1981) to argue that reactance was not immeasurable (Ratcliff, 2021; Rosenberg & Siegel, 2018). The intertwined model proposed by Dillard and Shen (2005), which models reactance as a latent factor consisting of anger and negative cognition, is the standard in the field (Rains, 2013; Ratcliff, 2021; Reynolds-Tylus, 2019).

However, questions concerning the operationalization and measurement of reactance persist and there are multiple areas in reactance research that suffer from conceptual overlap and a lack of clarity (Ratcliff, 2021). Within the "intervening psychological response" category, some researchers use negative cognitions only, some use affective reactions only, and some do not measure any intervening states at all, sticking to the original focus on antecedents and consequences (Ratcliff, 2021). Further, many studies exclude anger as a measure and do not set up their studies to assess boomerang attitude or behavior thus confounding reactance with other types of failed persuasion (Ratcliff, 2021).

To help add conceptual clarity into the temporal process of reactance and understand the elusive nature of the intervening response between freedom threat and freedom restoration, scholars have been calling for research that uses neuroscience to help triangulate key propositions of psychological reactance theory (PRT) (Rains, 2013; Ratcliff, 2021). Almost all

reactance research is based on self-reported instruments and structural equation modelling of cross-sectional data (Rains, 2013), and neuroscience could help establish the causal order of the process and establish that the experience of reactance is not an artifact of self-reported measures. And, though the consensus is that anger and negative cognitions constitute reactance, only anger is uniquely associated with reactance and not confounded with other types of failed persuasion. In particular, the emotional component to reactance (i.e. anger) is less well-suited to self-report. The strength of the neuroscience approach is that it offers a way to directly explore mechanisms like emotional processing and memory rather than relying on indirect indicators (Amodio, 2010; Ochsner & Lieberman, 2001), which depend on the ability of individuals to accurately introspect about their thought processes and choices that occurred during stimuli exposure (Lieberman, 2010).

An ideal context to study reactance in the brain is graphic health warning labels for cannabis products. Warning labels provide a practically relevant, real-world context in which domineering language, often associated with reactance (Quick et al., 2013; Reynolds-Tylus, 2019), is the prevailing style of communication.

The current study is an fMRI companion study to a population-based survey experiment studying responses to cannabis warning labels among adolescents and young adults by Yang et al. (2024). At present, this analysis is concerned with assessing whether the effect of persuasive advertisements for edible cannabis products are influenced by graphic health warning labels and whether the presence of warning labels increases neural activity in regions associated with anger (amygdala) and negative affect (dorsal anterior cingulate (dACC) and anterior insula (AI)) (Barrett & Bliss-Moreau, 2009; Richard et al., 2023), regions that may underlie the emotional component of reactance (Dillard & Shen, 2005, Ratcliff, 2021).

Graphic Health Warning Labels

The effort to evaluate and implement graphic health warning labels started with cigarettes, and now more than 100 countries mandate graphic health warning labels for cigarettes (Shadel et al., 2019). The U.S. Federal Drug Administration issued a rule to mandate graphic health warning labels in 2012, and a final rule in favor of graphic health warning labels for cigarettes in 2020 (Federal Drug Administration, 2020), but a long series of legal challenges have led to years of delays (Shadel et al., 2019; *Tobacco Reporter*, 2022). Therefore, the evidence regarding warning labels for tobacco products in the U.S. is limited to academic research tests.

To assess the impact of graphic health warning labels, Brewer and colleagues (2016) conducted an RCT comparing the status quo labels with the graphic warning labels proposed among a convenience sample of adult cigarette smokers from the general population. They found that after four weeks, smokers that were provided cigarettes with the graphic health warning labels were more likely to attempt to quit smoking during the trial than those provided with cigarettes labelled in the status quo (40% vs 34%; odds ratio [OR], 1.29; 95% CI, 1.09-1.54). To simulate the real world as much as possible, Shadel et al. (2019) set up an experiment at the point-of-sale between the standard FDA approved text warnings and graphic health warning labels using the RAND StoreLab (RSL), a fake convenience store where tobacco products are displayed in the traditional power wall arrangement. The main effects analysis showed that there was no difference in purchasing behavior based on warning label condition. However, follow-up analyses showed that nicotine dependence interacted with experimental conditions such that those who were less dependent were less likely to buy cigarettes in the graphic health warning label condition compared to the text-only warnings.

In a recent meta-analysis of experimental studies across countries, Noar et al. (2020) found that pictorial warning labels for cigarettes outperformed text-only labels on cognitive elaboration and elicitation of fear and negative affect, but not beliefs about perceived severity, likelihood of harm or experiential risk. Noar et al. (2020) noted that existing evidence associating pictorial warning labels with increased quit intentions has not explicated a causal mechanism. Neural studies could provide an answer. There is now extant literature showing that neural activity in response to health messages predicts behavior at a population level (Falk et al. 2010a; 2012). Most relevant to the present research, Riddle et al. (2016) used VMPFC activation in response to graphic health warning labels for cigarettes to predict smoking behavior among U.S. adult smokers.

Alongside potential to nudge individuals toward reducing consumption of harmful products, warning labels also have the potential to incite reactance. LaVoie and colleagues (2017) found that graphic health warning labels increased perceptions of freedom threat and source domineeringness and predicted experiencing state reactance.

Interestingly, the experience of reactance in response to warning labels does not necessarily preclude all persuasive effects. A randomized trial with US cigarette smokers conducted by Hall and colleagues (2018a) found that experiencing reactance only partially reduced support for warning label policies among smokers in the trial. However, the same trial showed that reactance also partially reduced quit intentions (Hall et al., 2018b). A more recent study found that graphic health warning labels for alcohol, sugar-sweetened beverages and tobacco incited more reactance, but were rated as more effective and reduced product appeal (Hall et al., 2020). However, exposure to graphic health warning labels in the study was not associated with support for mandating warning labels at the policy level. Alternatively, reactance

to warning labels may actually drive desired effects. One longitudinal study of US cigarette smokers showed that reactance predicted forgoing cigarettes to avoid graphic health warning labels (Cho et al., 2016).

The research on GHWLs specific to cannabis is very limited. Kim and colleagues (2022) recently tested the California status quo label, which uses a composite health warning and text only, against single-warning text and single-warning text with pictorial enhancements. They found that single-themed messages with larger font, against yellow background and pictorially enhanced warning labels performed better than the current standard in California on recall measures. They also found that pictorial enhanced warning labels performed better on perceived message effectiveness. More recently, Yang and colleagues (2024) conducted a multi-message effects experiment with a representative sample of 970 adolescents and 1776 young adults susceptible to cannabis use. They showed that pictorial graphic health warning labels reduced cannabis product appeal and intentions to use cannabis for both adolescents and young adults.

The empirical evidence on reactance and cannabis warning labels is scant for younger audiences. Only Yang and colleagues (2024) have examined cannabis in teens and young adults, and a review by Francis et al. (2019) on the impact of graphic health warning labels for tobacco products on youth and young adults only included one reactance study.

Acute and Chronic Risk

The most relevant evidence we have about designing messages for adolescents and young adults to prevent harms from cannabis comes from the mass communication literature on preventing tobacco uptake and encouraging tobacco cessation. A wealth of evidence was produced by "The Real Cost" campaign, which represented a moonshot effort to end smoking

and vaping in American youth that ultimately reached nearly 90% of American teens within the first campaign year (Santiago et al., 2019a). "The Real Cost" was estimated to have prevented 380,000–587,000 youths aged 11–19 years from taking up smoking between 2016 and 2019 (Duke et al., 2019).

Although the campaign was targeted at teens, especially those who were open to trying cigarettes or were already experimenting with them (Santiago et al., 2019b), there is evidence that the campaign reached and influenced young adults as well (Hall et al., 2019). When analyzing beliefs of teens susceptible to trying smoking and those currently smoking, Santiago et al. (2019b) found that these teens were more likely to believe that smoking had no short-term consequences and that they could not become addicted. They also found that these susceptible teens felt they could make their own decisions and wanted to understand 'why' smoking is unhealthy. In short, Santiago et al. (2019b) described the gap in education as:

> "These at-risk teens believe some of the long-term health consequences of smoking, but live in the moment—thinking the next cigarette will not cause harm and they will stop way before they start to see any serious health consequences." (S12).

In a secondary evaluation of campaign concept tests, youth responded most to defining addiction as a 'loss of control' and responded more to ads depicting "realistic, short-term consequences of addiction" and "responded favorably to the Hacked concept's message describing how nicotine changes the brain," (Roditis et al., 2019, p. S29).

Ultimately, although message appeals to acute versus chronic risks have been studied primarily in the context of cigarette smoking, there is reason to believe the same dynamics would hold for cannabis use.

Brain as Predictor

Many fMRI studies aim to "map" functions of the brain to specific structures and regions or areas, since brain function is somewhat localized (Poldrack, 2018). However, most areas of the brain are activated for multiple functions. Therefore, mapping of an area is not exclusive to any function. The fact that multiple processes activate the same regions is the fundamental problem with reverse inference (Poldrack, 2006). The standard method of inference in neuroimaging is to design a 'task' which is known to engage certain functions and then look at the pattern of activation.

Based on foundational studies that mapped activation of certain regions with certain functions, Berkman and Falk (2013) identified an approach to using individual differences in brain activation to predict outcomes of interest. The brain-as-predictor approach often captures variance in outcome measures that is unique from data extracted from traditional social science data, often self-report. This unique variance may be due to a variety of factors, such as the influence of social desirability, unconscious influences that participants are unaware of, or the added cognitive step of introspection during self-report (Lieberman, 2010). Moreover, the brain-as-predictor approach side-steps the problems with attempting to use neural signals alone as evidence that a psychological process occurred. The approach incorporates real-world data with neural data to test whether brain activation predicts behavioral outcomes (Falk et al., 2010a,b; 2012). The advantages of brain-as-predictor approaches are "the ability to predict variance that is unique to what is explained by certain self-report measures and evidence supporting links between key psychological mechanisms stimulated by message exposure (e.g., self-related processing and valuation) and prediction of key behavioral outcomes," (Falk et al., 2015, p. 40).

Thus, the current study aims to leverage this approach as a way of assessing the influence of warning labels on message processing to examine whether the inclusion of warning labels or counter attitudinal information may influence regions associated with behavior change in response to persuasive information (Bartra et al., 2013; Falk et al. 2010a, b; 2012) and whether warning labels incite activity in regions that may underlie the emotional component of reactance (Barrett & Bliss-Moreau, 2009; Dillard & Shen, 2005; Ratcliff, 2021; Richard et al., 2023).

Persuasion in the Brain

The fundamental process behind effective persuasion and behavior change, for persuaders and receivers, is subjective value maximation. Meaning, "people make choices to maximize the value they expect from their actions," (Falk & Scholz, 2018, p. 330). The major inputs to subjective valuation for persuaders and receivers are self-relevance and social relevance. Therefore, the mental processes that fMRI researchers use as indicators of persuasion are a) self-referencing, b) mentalizing (i.e. thinking about others), and c) valuation. This account suggests that during effective persuasion, individuals integrate information about the value placed on message content with respect to their own goals and motivations (Falk et al. 2010a, b; 2012). Beyond the self, thinking about how others think and feel (i.e. mentalizing), facilitates persuasion between a communicator and a receiver by driving both to think about the social consequences of the message (Baek et al., 2020), consistent with the overall idea that social relevance is an input to subjective valuation. Further, the role of mentalizing in persuasion has been validated cross-culturally and across mediums (Falk et al., 2010b). Finally, the inputs of self and others culminate in valuation activity, the "intermediate computational step during decision making, in which alternatives are placed on a common scale to facilitate value-maximizing

choice," (Bartra et al., 2013, p. 412). Further, the valuation network reflects "decision making and outcome delivery, as well as for both monetary and primary-reward outcomes," (p. 426).

Since the purpose of warning labels is to both inform consumers of potential harms of cannabis, and since research from Riddle et al. (2016) showed that VMPFC activity is associated with warning labels, the first hypothesis is advanced.

H1: Participants will exhibit less activity in the VMPFC during encoding of messages with a given warning label than during encoding of messages with no warning labels.

Anger and Negative Affect in Brain

Reactance is a motivational state that drives individuals to restore freedom in response to threats to freedom, such as restrictions or attempts at social influence. The two defining features of the "intervening psychological response" that underpin motivation in response to a freedom threat are negative cognition and anger (Rains, 2013; Ratcliff, 2021). Thus, the amalgamation of negatively valanced thoughts and anger are taken as primary indicators of reactance following any potential freedom threat. However, negative cognitions are not exclusive to reactance, but could indicate other message processes associated with other theories, like the elaboration likelihood model (Ratcliff, 2021).

Therefore, it makes sense to use anger in response to persuasive messages as a proxy indicator of reactance in a neuroscience study. In the reactance literature, reactance is discussed as having affective and cognitive components (Ratcliff, 2021). However, affect itself is a complex process of the mind that entails taking in physiological signals, interpreting and responding. The complicated relationship between cognition and affect holds whether the research uses a lens of discrete emotions (e.g. anger, sadness) or appetitive versus aversive

dimensions (Lee & Lang, 2009; Harmon-Jones et al., 2017). Due to the complex interweaving of cognition and arousal, functional mapping in neuroscience has failed to consistently find that discrete emotions are tied to distinct structures or regions in the brain (Minich et al., 2023b). As an alternative to the idea that the discrete emotions have different patterns of functional activation, some scholars have advanced the idea that emotions are cognitive interpretations of "a general neurophysiological state that shifts in response to an individual's environment" known as core affect, which has more empirical support (Barrett & Bliss-Moreau, 2009; Lindquist et al., 2012; Minich et al., 2023b). The appraisal theory of emotions and the discrete emotions perspective in communication hold that emotions also rely on a process of being aroused by an internal or external stimuli to then interpret and act upon in *distinct categories that predict outcomes* (Lazarus, 1991; Nabi, 2010). Early neural studies sought to find the same distinct categories from the discrete emotion perspective in the brain but have failed to do so and instead advanced models to measure and predict core affect.

Core affect may be activated in the "neuronal workspace" that includes the limbic system and frontal brain regions (Barrett & Bliss-Moreau, 2009; Minich et al., 2023b). The anterior insula, the amygdala and the dorsal anterior cingulate cortex (dACC) have an association with the visceromotor network within the affective workspace that drives autonomic processes in the body and, ultimately, autonomic behavior (Barrett & Bliss-Moreau, 2009; Minich, 2023 et al., 2023b). Further, the anterior insula, the amygdala and the (dACC) have been shown to predict anger and threat responses (Richard et al., 2023)

In a recent companion study looking at reactance in the brain, Minich (2023) drew a parallel between PRT and game theory studies in neuroscience showing that anterior insula activity is associated with punishing unfair players even if it goes against the best interest of the

individual. Minich (2023) also showed in a study of anti-driving under the influence of cannabis public service announcements that PSA ranking by self-reported anger corresponded with intersubject correlation of activity in the anterior insula.

The numerous studies showing that warning labels incite reactance across product categories (Cho et al., 2016; Hall et al., 2018a, 2018b, 2020) and the evidence linking activity in the anterior insula, amygdala, and dorsal anterior cingulate cortex (dACC) to core affect and anger (Barrett & Bliss-Moreau, 2009; Richard et al., 2023) justify the advancement of the following hypotheses.

H2: Participants will exhibit more activity in the anterior insula during encoding of messages with a given warning label that is a) acute and b) chronic than during encoding of messages with no warning label.

H3: Differences in anterior insula activation during encoding of warning labels compared to no warning labels will be positively associated with self-reported anger for a) acute and b) chronic graphic health warning label.

H4: Participants will exhibit more activity in the amygdala during encoding of messages with a given warning label that is a) acute and b) chronic than during encoding of messages with no warning label.

H5: Differences in amygdala activation during encoding of warning labels compared to no warning labels will be positively associated with self-reported anger for a) acute and b) chronic graphic health warning label.

H6: Participants will exhibit more activity in the dACC during encoding of messages with a given warning label that is a) acute and b) chronic than during encoding of messages with no warning label.

H7: Differences in dACC activation during encoding of warning labels compared to no warning labels will be positively associated with self-reported anger for a) acute and b) chronic graphic health warning label.

Methods

Consistent with the principles of open science (Spellman, et al., 2018), all materials needed to reproduce this study are publicly available on the Open Science Foundation website: https://osf.io/9kyhg/?view_only=7c20b75353364319977fab4426ed2945

Participants

Forty right-handed participants ($N = 40$) were recruited from a Midwestern college town using a university job board. Data was collected in the Summer and Fall of 2022. Eligible participants were adolescents 18-24 who scored at least one point ($M = 7.46$, $SD = 5.19$, range 1-19) on the Cannabis Use Disorder Identification Test (CUDIT) developed by Adamson and Sellman (2003). Participants were excluded for reasons that make the vulnerable to research exploitation, including being incarcerated or in a treatment facility, having impaired decision-making abilities, receiving treatment for mental illness, substance abuse or developmental disability, having served in the military, or indicating vulnerability. Participants were also excluded if they could not safely get an fMRI for reasons including suffering from claustrophobia, taking psychoactive drugs, having a history of psychiatric or neurological

disorders, having uncorrected vision or having any metal in their body that is contraindicated for fMRI. This study only included individuals who were right-handed.

Procedure

After obtaining consent, participants completed a series of pre-scan questionnaires (unrelated to this analysis), fMRI safety screening, and were trained on fMRI tasks. Participants then underwent a one-hour scan session that included several tasks, including a cannabis warning label (CWL) task. Immediately after the scan, participants completed another set of questionnaires that assessed multiple persuasion outcomes not analyzed here.

Cannabis Warning Label (CWL) Task

While undergoing fMRI, participants viewed a series of advertisement conditions in a rapid event-related design. All participants completed two consecutive eight-minute runs of fMRI, viewing 60 cannabis marketing posts. Advertising conditions were constructed for this experiment and varied across two fully crossed conditions: each used either a graphic health warning label or no warning label, presented alongside pro-cannabis, anti-cannabis, or filler text. Conditions were assigned using a randomized, counterbalanced design. Participants viewed each social media advertisement pairing for a period of seven seconds, then viewed a set of social media comments that were manipulated by valence (positive: negative: neutral). Participants viewed these comments for a period of six seconds, then responded to the prompt "How effective is this cannabis ad?" on a scale of 1 (not effective) to 4 (very effective) using a four-button Current Designs response pad held in the participant's right hand. Participants were given three seconds to respond to this prompt, then were presented with a fixation cross for a randomized, jittered period of an average of 1.5 seconds.

Stimuli

This study systematically varied stimuli to reduce case category-confounding (Jackson, 1992; Slater et al., 2015) and to ensure that effects were not specific to only one type of warning message, cannabis advertisement, or set of social media comments. First, following the protocol described in a published paper (Kim et al., 2022), ten single-themed CWLs with textual and visual enhancements were developed, each using a 12-point font and covering approximately 30% of the marking post (see Figure 1).

Each of these CWLs addresses one of the following health risks supported by the current state of cannabis science (National Academies of Sciences, Engineering, and Medicine et al., 2017; Volkow et al., 2016) and cited in a recent bill introducing graphic CWLs in California (An Act to Add Sections 26070.3 and 26121 to the Business and Professions Code, Relating to Cannabis, 2022): early use and cognitive function loss, driving risks, mental health issues, suicidal ideation, delayed effects, exposure to toxic contaminants, nausea and vomiting, and mother-to-baby transfer of substances. Further, a subset of warning labels was used to compare acute and chronic risk. See Table 1 for the content of warnings used in the study.

All posts were collected after January 2021. Marketing posts were vetted by the study team to ensure that they have employed youth-targeted appeals such as candy-like portrayals and fruity-flavor emphasis. Comments were primarily pulled from Reddit. Comments were pre-processed to remove references to specific original users and other posts, and then coded by the study team for valence, with a neutral condition made of filler text (See Figure 2). Lastly, each of the ten CWLs was imposed onto the set of original marketing posts.

Figure 1. Example of advertisements with acute and chronic graphic health warning labels.

fMRI Data Acquisition and Pre-Processing

Structural and functional brain imaging was conducted using a 3 Tesla GE Discovery MR750 scanner. Head motion was minimized using foam padding on the head coil. Two functional runs were recorded (TR = 800ms, TE = 20, flip angle = 60°, matrix size = 96x96, 54 axial slices, 3mm thick; voxel size = 3.0x3.0x3.0) for 36 participants, and one run was recorded for the remaining four participants. A motion-corrected T1-weighted MPnRAGE acquisition with 1.0 mm isotropic spatial resolution was used as an anatomical underlay (Kecskemeti et al., 2018). Image preprocessing and analysis were performed using the afni_proc.py program within the Analysis for Functional Neuroimaging (AFNI) software package (Cox, 1996). Functional and anatomical runs were warped to align with the Montreal Neurological Institute (MNI) template brain and smoothed with a 4mm Gaussian kernel. To ensure only steady-state images were used in our analysis, the first 7 TRs (5600 ms) were discarded.

The effects of warning labels were estimated with voxel-wise first-level regression models for message subsets of interest, specifically regarding acute and chronic risks of cannabis use. Each model includes messages as regressors, as well as regressors for motion, and nuisance regressors using AFNI's 3dDeconvolve command.

Figure 2. Examples of neutral, pro-cannabis, and anti-cannabis comment conditions
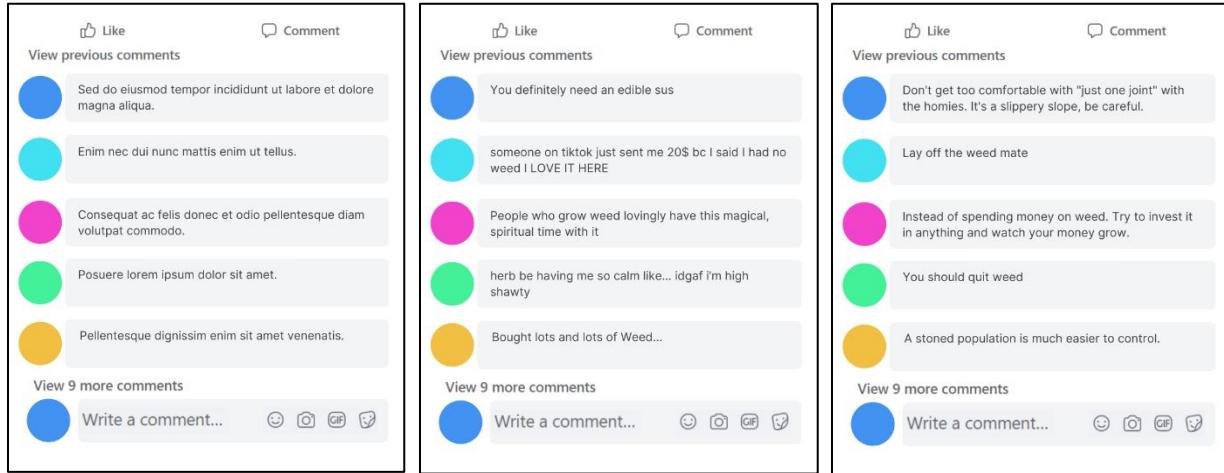
Post-fMRI Survey

After completing the fMRI tasks, participants included a post-survey. In the post-survey, participants viewed cannabis advertisements with a randomly selected a) acute and b) chronic health warning label and completed a series of reactance measures of each. This study used a self-reported measure of anger to see if neural activity predicted self-reported data.

Anger

Anger toward the warning label was assessed by asking participants how much participant felt irritated, angry, aggravated and, irritated on a scale from 0 = "None of this feeling" to 4 = "A great deal of this feeling." (Dillard & Shen, 2005). Respondents reported their

anger toward both a) an acute health warning label ($M = 1.53$, $SD = .72$) and b) a chronic health

warning label ($M = 1.49$, $SD = .79$).

Table 1

Content of warning labels and categorization scheme

Warning	Category
Buy Legal! Illegally sold cannabis is more likely to contain unsafe additives or harmful contaminants such as mold or pesticides.	Uncategorized
Do not use if pregnant or breastfeeding. Substances in cannabis are transferred from the mother to the child and may harm your baby's health, including causing low birth weight.	Chronic
Cannabis use may contribute to mental health problems including psychotic disorders, such as schizophrenia. Risk is greatest for frequent users, and with use of products high in THC.	Chronic
Cannabis use may contribute to mental health problems including increased thoughts of suicide and suicide attempts. Risk is greatest for frequent users.	Chronic
Driving while high is a DUI. Cannabis use increases your risk of motor vehicle crashes.	Acute
Not for Kids or Teens! Starting cannabis use young or using frequently may lead to problem use and, according to the U.S. Surgeon General, may harm the developing brain.	Chronic
Prolonged use of cannabis products high in THC may cause recurrent, severe nausea and vomiting.	Acute
It can take up to 4 hours to feel the full effects from eating or drinking cannabis. Consuming more within this time period can result in adverse effects that may require medical attention.	Acute
The higher the THC content, the more likely you are to experience adverse effects and impairment. THC may cause severe anxiety and disrupt memory and concentration.	Uncategorized
Co-use of cannabis and alcohol can increase your risk of motor vehicle crashes more than using cannabis or alcohol alone. Combining substances also increases the risk of alcohol poisoning and accidental injuries.	Acute

Note. All labels started with 'WARNING.'

Region of Interest Analyses

This analysis examines activation patterns associated with four *a priori* regions of the brain: the ventral medial prefrontal cortex (VMPFC), the anterior insula, the amygdala, and the dorsal anterior cingulate cortex (See figures 3 through 6). The neural mask for the anterior insula is replicated from Minich (2023) and the mask for the VMPFC was initially created by Vezich et al. (2017) and replicated by Minich et al. (2023a) and Minich (2023). Coordinates of the regions of interest are also presented in Table 2.

Table 2.

Coordinates for *a priori* regions of interest

	X	Y	Z	Size (Voxels)
VMPFC	0	60	-9	33
AI	-32.5	-10.8	-24.5	925
	35	-13.2	-24.5	771
Amygdala	-20	4.2	-27	87
	25	4.2	-27	90
dACC	7.5	-28.2	13	1384

Note. Coordinates are for the MNI template brain.

Figure 3. Ventral medial prefrontal cortex (VMPFC) region of interest
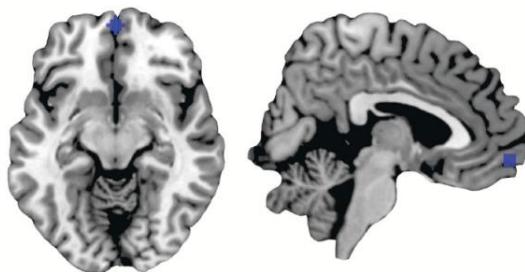
Note. Region of interest overlaid on the MNI template brain using MRICro.

Figure 4. Anterior insula (AI) regions of interest

Note. Region of interest overlaid on the MNI template brain using MRICro.

Figure 5. Amygdala regions of interest
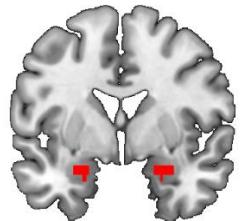
Note. Region of interest overlaid on the MNI template brain using MRICro.

Figure 6. Dorsal anterior cingulate cortex (dACC) region of interest.
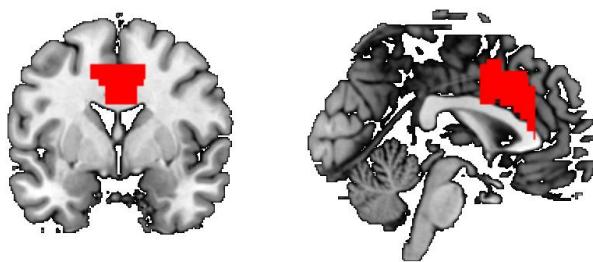
Note. Region of interest overlaid on the MNI template brain using MRICro.

Statistical Tests

Statistical tests were run on all participants who completed the warning label task and finished post-scan surveys. Three participants failed to complete the post-scan survey and thus were excluded ($N = 37$).

Tests of warning labels on neural activity

First, simple comparisons were conducted to determine whether warning labels influence neural activity. Specifically, one-sample t-tests comparing a) acute warning labels to no warning labels, and b) chronic warning labels to no warning label, for each of the regions of interest: ventral medial prefrontal cortex (H1), anterior insula (H2), amygdala (H4), and the dorsal anterior cingulate cortex (H6).

Using neural activity to predict self-reported anger

To test whether neural activity predicts self-reported anger, this study used linear mixed effects models for each region of interest associated with anger: anterior insula (H3), amygdala (H5) and dorsal anterior cingulate cortex (H7). For each region of interest associated with anger, one model was fitted using the contrast of a) acute risk warning labels compared to no warning labels and the other used the contrast of b) chronic risk warning labels compared to no warning labels. To account for the nested nature of our data, we also included by-participant and by-stimulus random intercepts and random slopes for all fixed effects.

Exploratory Whole-Brain Analyses

To test for effects of warning labels outside *a priori* networks of interest, an exploratory whole-brain analyses were conducted using the AFNI command 3dttest++ that contrasted the encoding of cannabis advertisements paired with a) acute health warning labels and b) chronic health warning labels compared to cannabis warning labels that were not paired warning labels. The result of this analysis was a whole-brain map in which each voxel contained a value representing the effect of our warning label condition on the neural signal in that voxel. Whole brain analyses used the voxel-level threshold of $p = 0.005$ and a cluster threshold of $p < 0.05$,

established through a Monte-Carlo-style simulation (10,000 iterations) using the AFNI command ClustSim.

<center>**Results**</center>

Region of Interest Analyses

Differences in activation for the *a priori* VMPFC ROI based on whether there was an acute warning label, or no warning label, were examined first. The comparison of acute warning labels to no warning labels showed that there was stronger VMPFC activity in the no warning label condition, $t(36) = -2.36$, $p = .02$. The comparison of chronic warning labels to no warning labels also showed that there was stronger VMPFC activity in the no warning label condition, $t(36) = -2.42$, $p = .02$. Thus, H1 was supported.

Differences in activation for the *a priori* anterior insula ROI based on whether there was an acute warning label, or no warning label, were examined first. The comparison of acute warning labels to no warning labels showed differences in activity in the anterior insula that could be due to sampling error, $t(36) = -1.79$, $p = .08$. The comparison of chronic warning labels to no warning labels showed anterior insula activation was stronger in the no warning label condition, $t(36) = -2.47$, $p = .02$. For both the acute and chronic health warning labels, the difference in anterior insula activation was greater in the no warning label condition, against expectation of H2.

Differences in activation within the anterior insula for acute health warning labels was not associated with self-reported anger toward the warning labels $\beta = -0.002$, $F(1, 35) = 0.002$, $p = 0.97$, nor was differences in activation within the anterior insula for chronic health warning labels, $\beta = 0.03$, $F(1, 35) = 0.97$, $p = .33$. H3 was not supported.

The comparison of acute warning labels to no warning labels showed no differences in activity in the amygdala, $t(36) = -0.54$, $p = .59$. The comparison of chronic warning labels to no warning labels also showed no differences in activity in the amygdala, $t(36) = -0.09$, $p = .93$. H4 was not supported.

Differences in activation within the amygdala for acute health warning labels was not associated with self-reported anger toward the warning labels $\beta = -0.06$, $F(1, 35) = 1.14$, $p = 0.29$, nor was differences in activation within the amygdala for chronic health warning labels, $\beta = -0.003$, $F(1, 35) = 0.003$, $p = .95$. H5 was not supported.

The comparison of acute warning labels to no warning labels showed no differences in activity in the dorsal anterior cingulate cortex, $t(36) = -1.03$, $p = .31$. The comparison of chronic warning labels to no warning labels also showed no differences in activity in the dorsal anterior cingulate cortex, $t(36) = -0.09$, $p = .31$. H6 was not supported. Differences in activation within the dorsal anterior cingulate cortex for acute health warning labels was not associated with self-reported anger toward the warning labels $\beta = -0.02$, $F(1, 35) = 0.17$, $p = 0.68$, nor was differences in activation within the dorsal anterior cingulate cortex for chronic health warning labels, $\beta = 0.03$, $F(1, 35) = 0.65$, $p = .43$. H7 was not supported.

Exploratory Whole-Brain Analyses

Exploratory whole-brain analyses failed to yield any distinct differences in patterns of activation between cannabis advertisements presented with acute health warning labels and advertisements presented without warning labels. They also failed to yield any distinct differences in patterns of activation between cannabis advertisements presented with chronic

health warning labels and advertisements presented without warning labels. Tables 3 and 4 present clusters that showed significant differences in activation for each contrast of interest.

Table 3

Clusters from Whole-Brain Analysis of Acute Health Warning Label > No Label Contrast

Region	X	Y	Z	Peak Z	Voxels
Cerebellum	-40.0	81.8	-19.5	-4.00	1898
Left Middle Occipital Gyrus	25.0	96.8	25.5	-2.89	428
Middle Temporal Gyrus	70.0	16.8	-4.5	3.02	410
Left Lingual Gyrus	0.0	69.2	5.5	3.39	231

Note. Clusters of significant neural activity associated with exposure to posts paired with warning labels referencing acute health risks as contrasted with exposure to posts not paired with warning labels. Whole brain results are corrected for multiple comparisons using a threshold of ($p = 0.005$, $K > 129$), corresponding to a corrected threshold of $p < 0.05$. "Peak Z" denotes the highest z-score within the cluster.

Table 4

Clusters from Whole-Brain Analysis of Chronic Health Warning Label > No Label Contrast

Region	X	Y	Z	Peak Z	Voxels
Right Cerebellum	-52.5	69.2	-19.5	-3.42	2108
Left Middle Temporal Gyrus	70.0	16.8	-7.0	2.86	826
Left Middle Occipital Gyrus	25.0	99.2	20.5	-2.81	319
Left Lingual Gyrus	0.0	71.8	5.5	4.39	226
Left Inferior Parietal Lobule	55.0	46.8	55.5	-3.55	152
Right Superior Occipital Gyrus	-22.5	91.8	35.5	-3.35	133
Left Superior Temporal Gyrus	57.5	-3.2	3.0	-3.13	100

Note. Clusters of significant neural activity associated with exposure to posts paired with warning labels referencing chronic health risks as contrasted with exposure to posts not paired with warning labels. Whole brain results are corrected for multiple comparisons using a threshold of ($p = 0.005$, $K > 95$), corresponding to a corrected threshold of $p < 0.05$. "Peak Z" denotes the highest z-score within the cluster.

Discussion

The goals of this study were to assess the ability of proposed cannabis warning labels to negatively influence the persuasive effects of social media advertisements for edible cannabis products. Additionally, this study was interested in whether warning labels would incite anger and negative affect, both indicators of reactance. This study also hoped to use neuroscience

methods to help triangulate self-reported measures of anger. Anger and negative affect were utilized as signs to attempt to distinguish potential reactance from other kinds of failed persuasion.

VMPFC activation was higher in the no warning label condition compared to acute warning labels and chronic warning labels. These results are consistent with findings from Riddle et al. (2016), in a tobacco context, that warning labels generated differences in VMPFC activity. This is also consistent with recent research from Minich et al. (under review) using the same stimuli which found that warning labels predicted differences in activation in a self-processing networking which overlapped with the VMPFC region of interest used in this study. Further, based on the general account of the valuation network activation as an indicator of persuasion and behavior change (Baek et al., 2020; Bartra et al., 2013; Falk et al. 2010a,b; 2012), these findings suggest that cannabis advertisements with warning labels are perceived differently than those without.

Both the comparison of acute and chronic health warning labels to no warning labels showed differences in activity in the anterior insula, but in the opposite direction that was hypothesized. Thus, against expectation, anterior insula activation was higher in the no warning label condition. Further, anterior insula activation when viewing acute health warning labels compared to no warning labels failed to predict self-reported anger. Anterior insula activation when viewing chronic health warning labels compared to no warning labels also failed to predict self-reported anger. Activity in the amygdala and dorsal anterior cingulate cortex, both regions of the brain associated with anger, did not differ between acute health warning labels compared to no warning labels, nor for chronic health warning labels compared to no warning labels. Further, differences in activity in the amygdala and dorsal anterior cingulate cortex failed to predict self-

reported anger. Exploratory whole-brain analyses failed to detect any distinctive activation patterns between acute and chronic health warnings associated with cannabis use when compared to cannabis advertisements without warning labels.

There are two possible explanations for the lack of activation of regions associated with anger in response to warning labels. The first is that reactance did occur, but this study failed to detect it in the brain because the experience of reactance has an alternative neural signature outside of anger and negative affect that we failed to test. The second is that reactance and anger did not occur in the participants in response to cannabis warning labels. The evidence for core affect being associated with activity in the anterior insula, amygdala and dorsal anterior cingulate cortex is quite strong (Barrett & Bliss-Moreau, 2009; Barrett, 2013), but the literature on reactance to warning labels among young people is quite scarce. Fundamentally, a lack of response in structures associated with core affect means that there is no evidence study participants experienced any affective reaction, thus no evidence they experienced anger or negative affect. It seems very plausible that individuals did not find the warnings particularly freedom threatening and thus never experienced any affective signs of reactance. Most reactance research in the field relies on researcher-contrived messages explicitly designed to induce reactance. It's possible that many common real world health messages will not incite reactance.

Practical Implications

This study has two important insights for the ongoing policy debate around cannabis warning labels. First, acute and chronic warning labels were associated with less VMPFC activation than cannabis ads without warning labels. Although both the ads for cannabis themselves and the warning labels are persuasive messages, it's reasonable to assume that young adults who use cannabis should have a positive valuation toward cannabis advertisements. Thus,

a difference between valuation scores for acute or chronic labels compared to ads with no warning labels such that a valuation signal that is lower is suggestive. Specifically, this could be interpreted as the warning label interfering with the persuasive intentions of the ad, as policymakers intend. The second insight from this study is that there was no evidence of brain activity associated with core affect predicting self-reported anger toward the warning label. Taken together, these findings may suggest that warning labels disrupt reward processing associated with cannabis product use, without any evidence that the warning labels are inciting an emotional or reactant response.

The findings from this study, alongside the findings from its companion study from Yang et al. (2024) which show that graphic warning labels reduced cannabis use intentions in teens and young adults, suggest that graphic health warning labels for cannabis products be enacted.

Limitations and Future Research

The results of this laboratory study should be considered with the inherent limitations of the study. This study relied on a crude measure of cannabis use. Future research could examine the effects of warning labels based on the amount of cannabis use or dependence. Common sense and formal health communication theory, like the Stages of Change model (Prochaska & DiClemente, 1983; DiClemente & Prochaska, 1998), recognizes that casual users or those interested in using are different from current users or even current heavy users. Further, Shadel et al. 2019 showed that cigarette warning labels influence purchasing behavior, but only for individuals lower in nicotine dependence. Additionally, the core strategy of the most successful campaign against cigarettes recognized that audience beliefs differ based on usage, and that messaging goes further for those not already using regularly (Santiago et al., 2019b). Future

cannabis warning label research should investigate cannabis use and dependence as moderators for the effects of warning labels.

Second, the current study was a secondary data analysis aiming to examine the underlying neural mechanism associated with emotional components of reactance. The current study design could be improved upon in future research that aims to examine these processes. A study which uses stimuli known to incite reactance (i.e. domineering language) and prompts participants about perceived freedom threat could elucidate the neural mechanisms of reactance more directly. Once a neural signature is more clearly identified and replicated, neural studies could help correct biases in self-reported measures, particularly related to the affective component of reactance.

This study uses a convenience sample maintained through a college town, so it is subject to WEIRD biases in the data (Henrich et al., 2010). So, although the sample fits the context for use, future studies should look at probability, representative samples in the U.S. and cross-culturally. This study also excluded participants typically identified as 'vulnerable' for purposes of research recruitment and excluded those taking regular psychoactive drugs due to fMRI technology constraints. However, many people using cannabis are also on other psychoactive drugs and may be the most at-risk for unintended consequences (Stuyt, 2018, Woodhouse, 2022). Future research should make an effort to study those that are acutely vulnerable to cannabis misuse. Lastly, the artificial environment of a laboratory study is fundamentally different from everyday experience and, thus, more behavioral research in the field is needed.

Conclusion

This study provides evidence that warning labels disrupt the persuasiveness of cannabis advertisements. Further, there was no evidence that neural activity associated with anger and negative affect across three different brain regions differed between cannabis advertisements with warning labels and without. Additionally, activation differences based on condition in regions associated with anger and negative affect failed to self-reported anger toward warning labels. Ultimately, these findings were consistent with warning labels mitigating the persuasiveness of cannabis advertisements. There was no evidence that warning labels generated affective responses consistent with reactance.

Conclusions, Implications and Future Directions

The past 50 years of research into psychological reactance theory (PRT) have generated many interesting insights in clinical psychology and persuasion research (Rosenberg & Siegel, 2018). In a post-COVID era, PRT has new relevance to help predict responses to both requests for individual behavior change (e.g. Knapp et al., 2021) and responses to policy mandates (e.g. Kriss et al., 2022; Rains et al., 2022). Further, some have pushed for new outcomes like willingness to share a message (Shoenberger et al., 2021) or avoiding a message (Clayton et al., in press). However, reactance research has yet to move beyond its founding and development in experimental psychology to take on effects beyond those that individuals can self-report, in isolation, and beyond researcher-contrived messaging designed to incite reactance.

The goal of this dissertation was to push psychological reactance methods and theory into new realms across three studies. Study 1 examined two trait-level characteristics thought to influence reactance in response to a mandate message based on real-world policy trends. Study 2 examined how reactant responses to public service announcements were influenced by group dynamics using focus groups, a rare methodology for reactance research. Finally, study 3 looked for neural signs of anger and negative affect in response to a potential new policy intervention: warning labels for cannabis products.

Dissertation Review

The results of the three studies are reviewed, followed by a broader discussion about the theoretical, methodological and practical implications of the findings. This dissertation concludes with limitations of the study methodologies and calls for future research.

Individual Differences and Psychological Reactance in Response to Policy Mandates

The first study was a secondary analysis of a message effects experiment conducted by Kriss et al. (2022) seeking to examine moderation effects of two personality traits: reactance proneness and locus of control. Specifically, it was hypothesized that reactance proneness would moderate the relationship between perceived freedom threat and experienced reactance (i.e. state reactance). Furthermore, it was predicted that greater externalized locus of control would moderate the relationship between reactance and attitude. Neither reactance proneness nor locus of control were found to be moderators of effects, instead they both emerged as independent predictors. Reactance proneness predicted reactance such that those that measure higher on reactance proneness report higher scores on a measure of state reactance in response to a message about a campus COVID-19 vaccine mandate. Locus of control predicted attitudes toward a COVID-19 vaccine mandate such that those higher in an external locus of control showed more positive attitudes toward the vaccine mandate when controlling for reactance proneness, reactance, and freedom threat. Thus, despite not moderating the respective hypothesized pathways, there was still evidence for the theorized effects of both reactance proneness and locus of control using stimuli and a sample that was high in ecological validity. A series of interaction analyses with properties of threat failed to yield a clear picture, but nonetheless showed that individual differences and properties of threat interact in complicated ways. Thus, the evidence suggests that reactance proneness and locus of control are predictors that should be included in future reactance models to increase precision of estimates and are worth considering in furthering theoretical development on the ways that individuals differ. Further, knowing these relationships could support public health and policy advertising tailoring by psychographics.

Reactance and Persuasion in Social Settings

Study 2 examined reactance responses to public service announcements against driving under the influence of cannabis in a co-viewing setting. This study applied thought-listing and reactance coding of negative cognition and anger to focus group data. The study also explored signs of perceived freedom threat. In an innovation from most focus groups used in persuasion research, this study went beyond aggregating individual cognitions and instead examined group effects (Cyr, 2016; Hollander, 2004). There was a surprising lack of anger and perceived freedom threat.

As expected, groups varied in the proportion of positive, negative, and neutral sentiments expressed. Some variation was explained by differences in group make-up. Groups that were screened for cannabis before participating in the research were aware that the study was made up exclusively of cannabis users, but the groups without screening criteria did not know the behavioral norm of the focus group they were in. Groups that were screened for cannabis use expressed different proportions of sentiments than did the groups who were not screened for cannabis use. Groups, screened for cannabis use and unscreened, generated a similar proportion of positive thoughts. However, the groups screened for cannabis use generated a lower proportion of negative thoughts and a greater proportion of neutral thoughts than did the unscreened groups. There are numerous possible explanations for the results, and it is beyond the current method to determine which is driving the effect. Possibilities include that a) PSAs were designed to appeal to cannabis users, thus they were less critical of them, b) that non-users were biased by third person effects trying to think about what cannabis users would think, c) users and non-users were biased by social desirability in groups, d) or experience of being part of an ingroup (i.e. exclusively cannabis users) reduced reactance, a possibility suggested by

Graupmann (2012). Future research should examine the differences between general audiences and audiences made up of cannabis users exclusively, while also controlling the knowledge participants have about the cannabis use of co-participants. The combination of group makes up and knowledge about group make-up by individuals could help determine what drives group effects in persuasion. Further, Study 2 highlights that social context is an important factor for reactance theory to consider going forward.

Examining Neural Responses to Cannabis Warning Labels

Study 3 was an fMRI companion project to a population-based survey experiment testing cannabis warning labels among adolescents and young adults by Yang et al. (2024). The third study was interested in whether the persuasive effects of cannabis advertisements were influenced by graphic health warning labels and whether warning labels generate signals associated with psychological reactance. As predicted, cannabis advertisements with acute and chronic health warning labels showed less neural activity associated with positive valuation when compared to advertisements without warning labels, specifically activation in the ventral medial prefrontal cortex. This study also predicted that acute and chronic warning labels would be associated with greater activation in neural regions associated with core affect, particularly anger and negative affect. Therefore, study 3 examined activation in the anterior insula, amygdala and dorsal anterior cingulate cortex. Against expectation, anterior insula activation was higher for advertisements with no warning labels when compared to both acute and chronic health warning labels. Also, against expectation, anterior insula activation failed to predict self-reported anger. Activation in the amygdala and dorsal anterior cingulate cortex did not different between advertisements with and without warning labels, and activation in those regions failed to predict self-reported anger toward warning labels. Considering this was an exploratory, initial

study of the neural processes associated with reactance and that the findings were null, more research examining reactance from a neuroscience perspective is warranted.

Discussion

One goal of this dissertation was to think more concretely about individual and group effects. Study 1 showed that individual difference variables predict perceived freedom threat and freedom restoration processes to improve the estimate of effects. Further, individual differences interacted with threat magnitude and threat direction, both properties of freedom threats. Understanding the differences of individuals is a necessary precursor to understanding the importance of examining group effects. Individuals vary in personality but then are also subjected to social influence in co-viewing settings, as shown in Study 2. Groups varied widely in the proportion of sentiments, some of which was surely driven by personalities interacting in an unpredictable fashion, much as how the real-world media ecosystem functions in an era of co-viewing. Further, the differences among groups were partially explained by differences in group make-up. Groups that were made up of exclusively cannabis users generated a lower proportion of negative thoughts than did groups that were not screened for cannabis use. There is reason to believe the difference may have been related to the difference in knowledge of the behavioral norm of the group. The focus groups made exclusively of cannabis users also knew every other person used cannabis, whereas unscreened groups were trying to intuit the norms of the group and act accordingly. Study 2 demonstrates that reactance is a social phenomenon and must be studied as such. Studies focused exclusively on individual exposure to messages and individual-level analysis of outcomes cannot help us understand important societal level message effects, a topic which needs more attention in communication research (Cho & Salmon, 2007) and in reactance.

The second biggest insight across all three studies is that real-world health policy and persuasion messages likely incite much less reactance than the estimates shown in most of the literature, likely due to overreliance on studies which use researcher-derived and tightly controlled stimuli. Study 2 showed that focus group participants expressed tiny proportions of angry sentiment and no perceived freedom threat toward PSAs against driving under the influence of cannabis, across fear and guilt appeals. Further, some suggested the arguments were not threatening enough. The disparity of findings between typical reactance studies and study 2 may be related to the fact that PSAs are a richer media than most print-only message conditions used in reactance research, many of which are text-only. In study 3, cannabis warning labels also produced a surprising lack of differences in neural activation associated with anger and negative affect between viewing acute or chronic health warning labels compared to control. Further, neural activation associated with anger and negative affect did not predict self-reported anger, which should raise concerns about self-reported measures being upwardly biased due to artifacts of measurement or participant demand effects.

Future Directions

The three studies in this dissertation offer specific insights for reactance theory and for the applied health contexts that they were examined in. The three studies also share some broad methodological strengths and weaknesses which reflect challenges and opportunities in the broader mission of understanding the influence of mass mediated persuasion and reactance.

Mitigating Case-Category Confounding in Messages Studies

Message studies with multiple messages per variable feature of interest are essential to building stronger evidence on reactance. The biggest message features meta-analysis to date by

O'Keefe and Hoeken (2021) could speak somewhat to failed persuasion and to boomerang attitude change, but ideally a future large scale meta-analysis of outcomes specific to reactance, like anger and message derogation, is needed. The reason that multi-message conditions are so essential is because more realistic stimuli contain many features, even if a particular study is only interested in one. Slater and colleagues (2015) distinguished message variability and message heterogeneity. They define message variability as "the defined and operationalized features of messages in a given study," and message heterogeneity as "all message features that are undefined and unmeasured in a given study," (p. 3). O'Keefe (2015) and Slater et al. (2015) agree that whenever possible, messages should be randomly sampled from large populations of potential messages to mitigate the case-category confound in message effects as articulated by Jackson (1992). Some examples of study designs to mitigate case-category confounds come from Yang et al. (2018, 2024).

Reactance in the Field

Reactance has been driven by its founding in traditional experimental psychology and later adoption by clinical psychology (Rosenberg & Siegel, 2018). Thus, reactance research has focused on laboratory studies. The next discipline to pick up reactance theory was health communication, which also relies heavily on forced exposure designs that are cross-sectional (Slater, 2004; Jerit et al., 2013). The over-reliance on forced exposure designs that rely on solitary viewing and reporting makes processes of social diffusion, conformity, and other social effects difficult to anticipate based on the current reactance literature. Also, questions about exposure, attention, and memory in real life remain a blind spot for reactance.

Another approach is to move reactance beyond message centric studies and instead look at larger societal trends. For example, Fairchild et al. (2015; 2018) used a historical lens to make

predictions about when fear appeals would be effective, and when they would incite backlash. In the COVID era, time series analyses have been used to illuminate some trends in real world reactance. Rains and colleagues (2022) showed that political ideology and perceived infection risk influenced (non)compliance with mask mandates and forgoing masks after mandates ended. Most recently, Rains and Richards (2024) showed that state vaccine mandates for COVID-19 did not reduce COVID-19 vaccine uptake in the short term, but states that implemented COVID-19 vaccine mandates had lower uptake of COVID-19 boosters and flu vaccines in the longer term. This pattern would be consistent with indirect freedom restoration. Further, it illustrates the importance of research outside the lab over a longer time frame and the importance of thinking through multiple outcomes.

Population-Based Samples

The human samples in psychology, particularly those in experimental psychology, are biased toward WEIRD populations (Henrich et al., 2010). Reactance research suffers from an over-reliance on college student samples (Reynolds-Tylus, 2019). The over-reliance on college students is particularly problematic in the current U.S. context because the differences in public attitudes across many issues and values varies widely based on education (Pew Research Center, 2016; Haidt, 2007; 2012). Population-based survey experiments (Mutz, 2011) will be needed if reactance research ever hopes to inform cultural and political debates beyond college campuses.

Removing Normative Constraints

Most limitations of reactance research apply to the communication field more broadly. However, somewhat uniquely, reactance research is also hindered by normative constraints. In the contemporary wave driven by health communication and clinical psychology (Rosenberg &

Sigel, 2018), studies operate under the assumption that reactance must be mitigated, which stifles the possibility for new message comparisons and outcomes. The vast majority of studies examine reactance outcomes to health messages using a variety of persuasive messages features, but they never manipulate message features alongside the *stance* of the message. Therefore, it is impossible to know whether the same techniques work for both sides of a health issue, or if they only work in the affirmative. Recently, Quick et al. (2024) examined restoration postscripts, social norms, and message stance. This was the first time the stance was manipulated in a reactance study. Interestingly, the difference in perceived freedom threat between the message with a restoration postscript and without a restoration postscript was much higher in the condition that was also promotionally framed compared to the message that was framed as oppositional.

The assumption that reactance is always unwarranted and undesirable, and studies that only examine one stance of an issue, hinders the utility of reactance theory to explain human behavior. Fundamentally, public health efforts like needle exchanges, vaccine mandates, masking, age restrictions for tobacco, substance bans, and other health policies and standards of behavior also contain moral arguments (Humphreys & Piot, 2012). Thus, health policy and health behavior will be interpreted differently in a morally pluralistic world (Haidt, 2012; Atari et al., 2023). Freedom itself is a political and moral concept that varies at the individual level, as shown in Study 1. Reactance theory could be strengthened by investigating the interactions between freedom, freedom threats and morals across the U.S. and the world.

Another problem with the fixation on mitigation has left the possibility of inciting reactance as a persuasive strategy unexplored (Quick et al., 2013; Reynolds-Tylus, 2019). As reactance scholars have pointed out, one of the most successful anti-tobacco campaigns was an

appeal to reactance (Quick et al., 2013; Reynolds-Tylus, 2019). Specifically, as described by Reynolds-Tylus (2019) "the truth® campaign focuses on painting 'big tobacco' as a manipulative entity attempting to circumvent teenagers' choices and freedoms through lies and deception," (p. 10). Knowing when reactance is a barrier persuasion, or when reactance can be harnessed as a tool for (counter-)persuasion, is not possible until message effects remove normative constraints in their methods.

Conclusion

Across three studies with three different methods, this dissertation showed that reactance is influenced by individual personalities, specific messages and by group dynamics. Further, a focus on real-world stimuli shows that reactance indeed occurs, but perhaps to fewer kinds of health messages than studies using artificial messages may suggest. Future reactance research should further explore the differences between self-reported and objective measures of reactance and expand analyses beyond the individual to tackle group and societal effects.

References

Adamson, S. J., & Sellman, J. D. (2003). A prototype screening instrument for cannabis use

disorder: The Cannabis Use Disorders Identification Test (CUDIT) in an alcohol-dependent

clinical sample. *Drug and Alcohol Review*, *22*(3), 309–315.

https://doi.org/10.1080/0959523031000154454

Albarracín, D., Jung, H., Song, W., Tan, A., & Fishman, J. (2021). Rather than inducing

psychological reactance, requiring vaccination strengthens intentions to vaccinate in US

populations. *Scientific Reports*, *11*(1), 20796. https://doi.org/10.1038/s41598-021-00256-z

Amodio, D. M. (2010). Can neuroscience advance social psychological theory? Social

neuroscience for the behavioral social psychologist. *Social Cognition*, *28*(6), 695–716.

https://doi.org/10.1521/soco.2010.28.6.695

An act to add Sections 26070.3 and 26121 to the Business and Professions Code, relating to

cannabis, no. CA SB1097, 2021–2022 Regular Session (2022).

https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=202120220SB1097

Andor, M., Cox, J., Gerster, A., Price, M., Sommer, S., & Tomberg, L. (2022). *Locus of control

and prosocial behavior* (30359). https://doi.org/10.3386/w30359

Andreoli, V. A., Worchel, S., & Folger, R. (1974). Implied threat to behavioral freedom. *Journal

of Personality and Social Psychology*, *30*(6), 765–771. https://doi.org/10.1037/h0037529

Ashworth, B. (2020, September 29). How to host a virtual watch party. *Wired*.

https://www.wired.com/story/how-to-host-a-virtual-watch-party/

Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., & Dehghani, M. (2023). Morality

beyond the WEIRD: How the nomological network of morality varies across cultures.

Journal of Personality and Social Psychology, *125*(5), 1157–1188.

https://doi.org/10.1037/pspp0000470

Baek, E. C., Scholz, C., & Falk, E. B. (2020). The neuroscience of persuasion and information propagation: The key role of the mentalizing system. In K. Floyd & R. Weber (Eds.), *The handbook of communication science and biology* (pp. 122–133).

Ball, H., & Wozniak, T. R. (2022). Why do some Americans resist COVID-19 prevention behavior? An analysis of issue importance, message fatigue, and reactance regarding COVID-19 messaging. *Health Communication*, *37*(14), 1812–1819. https://doi.org/10.1080/10410236.2021.1920717

Ball, H., Weber, K., Goodboy, A. K., Kunkle, C. E., Lilly, C. L., & Myers, S. A. (2023). A mixed methodological examination of older adults' psychological reactance toward caregiving messages from their adult children. *Communication Monographs*, *90*(2), 137–158. https://doi.org/10.1080/03637751.2022.2128197

Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, *5*(4), 379–389. https://doi.org/10.1177/1754073913489753

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167–218. https://doi.org/10.1016/S0065-2601(08)00404-8

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Berkman, E. T., & Falk, E. B. (2013). Beyond brain mapping: Using neural measures to predict real-world outcomes. *Current Directions in Psychological Science*, *22*(1), 45–50. https://doi.org/10.1177/0963721412469394

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. Academic Press.

Brewer, N. T., Hall, M. G., Noar, S. M., Parada, H., Stein-Seroussi, A., Bach, L. E., Hanley, S., & Ribisl, K. M. (2016). Effect of pictorial cigarette pack warnings on changes in smoking behavior a randomized clinical trial. *JAMA Internal Medicine*, *176*(7), 905–912. https://doi.org/10.1001/jamainternmed.2016.2621

Cappella, J. N. (2006). Integrating message effects and behavior change theories: Organizing comments and unanswered questions. *Journal of Communication*, *56*(1), S265–S279. https://doi.org/10.1111/j.1460-2466.2006.00293.x

Cappella, J. N., & Rimer, B. K. (Eds.). (2006). Integrating theories of message effects and health behavior change to improve cancer control. *Journal of Communication*, *56*(S1), S1–S279.

Cheng, C., Cheung, M. W.-L., & Lo, B. C. Y. (2016). Relationship of health locus of control with specific health behaviours and global health appraisal: A meta-analysis and effects of moderators. *Health Psychology Review*, *10*(4), 460–477. https://doi.org/10.1080/17437199.2016.1219672

Cherulnik, P. D., & Citrin, M. M. (1974). Individual difference in psychological reactance: The interaction between locus of control and mode of elimination of freedom. *Journal of Personality and Social Psychology*, *29*(3), 398–404. https://doi.org/10.1037/h0035906

Cho, H., & Salmon, C. T. (2007). Unintended effects of health communication campaigns. *Journal of Communication*, *57*(2), 293–317. https://doi.org/10.1111/j.1460-2466.2007.00344.x

Cho, Y. J., Thrasher, J. F., Swayampakala, K., Yong, H.-H., McKeever, R., Hammond, D., Anshari, D., Cummings, K. M., & Borland, R. (2016). Does reactance against cigarette

warning labels matter? Warning label responses and downstream smoking cessation amongst adult smokers in Australia, Canada, Mexico and the United States. *PLOS ONE*, *11*(7), e0159245. https://doi.org/10.1371/journal.pone.0159245

Chung, A., & Rimal, R. N. (2016). Social norms: A review. *Review of Communication Research*, *4*, 1–28. https://doi.org/10.12840/issn.2255-4165.2016.04.01.008

Clayton, R. B., Reynolds-Tylus, T., Gonzalez, A. M., & Park, J. (in press). Beyond counterarguing? Comparing multiple measures of cognitive resistance for psychological reactance using the Avoidance, Contesting, and Empowering (ACE) resistance typology. *Motivation Science*. https://doi.org/http://dx.doi.org/10.1037/mot0000326

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Cyr, J. (2016). The pitfalls and promise of focus groups as a data collection method. *Sociological Methods and Research*, *45*(2), 231–259. https://doi.org/10.1177/0049124115570065

David, C., Cappella, J. N., & Fishbein, M. (2006). The social diffusion of influence among adolescents: Group interaction in a chat room environment about antidrug advertisements. *Communication Theory*, *16*(1), 118–140. https://doi.org/10.1111/j.1468-2885.2006.00008.x

De las Cuevas, C. (2023). Psychiatric patients' perceived health control and reactance: Implications for medication adherence. *Patient Preference and Adherence*, *17*, 1591–1601. https://doi.org/10.2147/PPA.S417608

De las Cuevas, C., Peñate, W., & Sanz, E. J. (2014). The relationship of psychological reactance, health locus of control and sense of self-efficacy with adherence to treatment in psychiatric

outpatients with depression. *BMC Psychiatry*, *14*(1), 324. https://doi.org/10.1186/s12888-014-0324-6

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (pp. 3–24). Plenum Press. https://doi.org/https://doi.org/10.1007/978-1-4899-1934-2_1

Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, *72*(2), 144–168. https://doi.org/10.1080/03637750500111815

Dillard, J. P., Tian, X., Cruz, S. M., Smith, R. A., & Shen, L. (2023). Persuasive messages, social norms, and reactance: A study of masking behavior during a COVID-19 campus health campaign. *Health Communication*, *38*(7), 1338–1348. https://doi.org/10.1080/10410236.2021.2007579

Duke, J. C., MacMonegle, A. J., Nonnemaker, J. M., Farrelly, M. C., Delahanty, J. C., Zhao, X., Smith, A. A., Rao, P., & Allen, J. A. (2019). Impact of The Real Cost media campaign on youth smoking initiation. *American Journal of Preventive Medicine*, *57*(5), 645–651. https://doi.org/10.1016/j.amepre.2019.06.011

Fairchild, A. L., Bayer, R., & Colgrove, J. (2015). Risky business: New York city's experience with fear-based public health campaigns. *Health Affairs*, *34*(5), 844–851. https://doi.org/10.1377/hlthaff.2014.1236

Fairchild, A. L., Bayer, R., Green, S. H., Colgrove, J., Kilgore, E., Sweeney, M., & Varma, J. K. (2018). The two faces of fear: A history of hard-hitting public health campaigns against

tobacco and AIDS. *American Journal of Public Health*, *108*(9), 1180–1186.

https://doi.org/10.2105/AJPH.2018.304516

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From neural responses to population

behavior: Neural focus group predicts population-level media effects. *Psychological

Science*, *23*(5), 439–445. https://doi.org/10.1177/0956797611434964

Falk, E. B., Berkman, E. T., Mann, T., Harrison, B., & Lieberman, M. D. (2010a). Predicting

persuasion-induced behavior change from the brain. *Journal of Neuroscience*, *30*(25),

8421–8424. https://doi.org/10.1523/JNEUROSCI.0063-10.2010

Falk, E. B., Cascio, C. N., & Coronel, J. C. (2015). Neural prediction of communication-relevant

outcomes. *Communication Methods and Measures*, *9*(1–2), 30–54.

https://doi.org/10.1080/19312458.2014.999750

Falk, E. B., Rameson, L., Berkman, E. T., Liao, B., Kang, Y., Inagaki, T. K., & Lieberman, M. D.

(2010b). The neural correlates of persuasion: A common network across cultures and

media. *Journal of Cognitive Neuroscience*, *22*(11), 2447–2459.

https://doi.org/10.1162/jocn.2009.21363

Falk, E., & Scholz, C. (2018). Persuasion, influence, and value: Perspectives from

communication and social neuroscience. *Annual Review of Psychology*, *69*(1), 329–356.

https://doi.org/10.1146/annurev-psych-122216-011821

Fawole, O. A., Srivastava, T., Fasano, C., & Feemster, K. A. (2018). Evaluating variability in

immunization requirements and policy among U.S. colleges and universities. *Journal of

Adolescent Health*, *63*(3), 286–292. https://doi.org/10.1016/j.jadohealth.2018.06.013

Federal Drug Administration. (2020). *Tobacco products; required warnings for cigarette

packages and advertisements (final rule) regulatory impact analysis*.

https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/tobacco-products-required-warnings-cigarette-packages-and-advertisements-final-rule-regulatory

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Psychology Press. https://doi.org/10.4324/9780203838020

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). McGraw-Hill.

Francis, D. B., Zelaya, C. M., Fortune, D. A., & Noar, S. M. (2021). Black college women's interpersonal communication in response to a sexual health intervention: A mixed methods study. *Health Communication*, *36*(2), 217–225. https://doi.org/10.1080/10410236.2019.1673949

Francis, D., Mason, N., Ross, J., & Noar, S. (2019). Impact of tobacco-pack pictorial warnings on youth and young adults: A systematic review of experimental studies. *Tobacco Induced Diseases*, *17*(May). https://doi.org/10.18332/tid/108614

Graupmann, V., Jonas, E., Meier, E., Hawelka, S., & Aichhorn, M. (2012). Reactance, the self, and its group: When threats to freedom come from the ingroup versus the outgroup. *European Journal of Social Psychology*, *42*(2), 164–173. https://doi.org/10.1002/ejsp.857

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*(5827), 998–1002. https://doi.org/10.1126/science.1137651

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion* (Kindle). Pantheon Books.

Hall, M. G., Grummon, A. H., Lazard, A. J., Maynard, O. M., & Taillie, L. S. (2020). Reactions to graphic and text health warnings for cigarettes, sugar-sweetened beverages, and alcohol: An online randomized experiment of US adults. *Preventive Medicine*, *137*(May), 106120. https://doi.org/10.1016/j.ypmed.2020.106120

Hall, M. G., Marteau, T. M., Sunstein, C. R., Ribisl, K. M., Noar, S. M., Orlan, E. N., & Brewer, N. T. (2018a). Public support for pictorial warnings on cigarette packs: An experimental study of US smokers. *Journal of Behavioral Medicine*, *41*(3), 398–405. https://doi.org/10.1007/s10865-018-9910-2

Hall, M. G., Saffer, A. J., & Noar, S. M. (2019). A secondary audience's reactions to "The Real Cost" advertisements: Results from a study of U.S. young adult smokers and susceptible nonsmokers. *American Journal of Preventive Medicine*, *56*(2), S57–S64. https://doi.org/10.1016/j.amepre.2018.08.005

Hall, M. G., Sheeran, P., Noar, S. M., Boynton, M. H., Ribisl, K. M., Parada Jr, H., Johnson, T. O., & Brewer, N. T. (2018b). Negative affect, message reactance and perceived risk: How do pictorial cigarette pack warnings change quit intentions? *Tobacco Control*, *27*(e2), e136–e142. https://doi.org/10.1136/tobaccocontrol-2017-053972

Harmon-Jones, E., Harmon-Jones, C., & Summerell, E. (2017). On the importance of both dimensional and discrete models of emotion. *Behavioral Sciences*, *7*(4), 66. https://doi.org/10.3390/bs7040066

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

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. https://doi.org/10.1080/19312450709336664

Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research*, *45*(4), 627–660. https://doi.org/10.1080/00273171.2010.498290

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hollander, J. A. (2004). The social contexts of focus groups. In *Journal of Contemporary Ethnography* (Vol. 33, Issue 5). https://doi.org/10.1177/0891241604266988

Hong, S.-M., & Faedda, S. (1996). Refinement of the Hong Psychological Reactance Scale. *Educational and Psychological Measurement*, *56*(1), 173–182. https://doi.org/10.1177/0013164496056001014

Horner, D. E., Sielaff, A., Pyszczynski, T., & Greenberg, J. (2023). The role of perceived level of threat, reactance proneness, political orientation, and coronavirus salience on health behavior intentions. *Psychology & Health*, *38*(5), 647–666. https://doi.org/10.1080/08870446.2021.1982940

Hornik, R., & Yanovitzky, I. (2003). Using theory to design evaluations of communication campaigns: The case of the National Youth Anti-Drug Media Campaign. *Communication Theory*, *13*(2), 204–224. https://doi.org/10.1111/j.1468-2885.2003.tb00289.x

Hornik, R., Jacobsohn, L., Orwin, R., Piesse, A., & Kalton, G. (2008). Effects of the National Youth Anti-Drug Media Campaign on youths. *American Journal of Public Health*, *98*(12), 2229–2236. https://doi.org/10.2105/AJPH.2007.125849

Huang, L., Huschka, T. R., Barwise, A. K., Allen, J.-S. P., Wolfersteig, W., Hamm, K., Cardenas, L. D., Phelan, S. M., & Allyse, M. A. (2024). Psychological reactance, misinformation, and distrust: A mixed methods analysis of COVID-19 vaccine uptake. *Journal of Clinical and Translational Science*, *8*(1), e48. https://doi.org/10.1017/cts.2024.15

Humphreys, K., & Piot, P. (2012). Scientific evidence alone is not sufficient basis for health

policy. *BMJ*, *344*(7854). https://doi.org/10.1136/bmj.e1316

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8),

0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Jackson, S. A. (1992). Empirical claims about message effects. In *Message effects research:*

Principles of design and analysis (pp. 1–23). Guilford.

Jeong, M., & Bae, R. E. (2018). The effect of campaign-generated interpersonal communication

on campaign-targeted health outcomes: A meta-analysis. *Health Communication*, *33*(8),

988–1003. https://doi.org/10.1080/10410236.2017.1331184

Jerit, J., Barabas, J., & Clifford, S. (2013). Comparing contemporaneous laboratory and field

experiments on media effects. *Public Opinion Quarterly*, *77*(1), 256–282.

https://doi.org/10.1093/poq/nft005

Judd, C. M., Yzerbyt, V.Y., & Muller, D. (2014). Mediation and Moderation. In H. Reis & C. M.

Judd (Eds.), *Handbook of Research Methods in Personality and Social Psychology* (2nd ed.,

pp. 653–676). Cambridge University Press.

Kecskemeti, S., Samsonov, A., Velikina, J., Field, A. S., Turski, P., Rowley, H., Lainhart, J. E.,

& Alexander, A. L. (2018). Robust motion correction strategy for structural MRI in

unsedated children demonstrated with three-dimensional radial MPnRAGE. *Radiology*,

289(2), 509–516. https://doi.org/10.1148/radiol.2018180180

Kim, J., Merrill, K., & Collins, C. (2020). Touchdown together: Social TV viewing and social

presence in a physical co-viewing context. *The Social Science Journal*, 1–15.

https://doi.org/10.1080/03623319.2020.1833149

Kim, S. J., Minich, M., Tveleneva, A., Liu, J., Padon, A. A., Silver, L. D., & Yang, S. (2022). Textual and pictorial enhancement of cannabis warning labels: An online experiment among at-risk U.S. young adults. *Drug and Alcohol Dependence*, *237*, 109520. https://doi.org/10.1016/j.drugalcdep.2022.109520

Knapp, M. E., Partington, L. C., Hodge, R. T., Ugarte, E., & Hastings, P. D. (2021). We're all in this together: Focus on community attenuates effects of pandemic-related financial hardship on reactance to COVID-19 public health regulations. *PLOS ONE*, *16*(12), e0260782. https://doi.org/https://doi.org/10.1371/journal.pone.0260782.g002

Krippendorff, K. (1970). Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement*, *30*, 61–70. https://doi.org/https://doi.org/10.1177/001316447003000105

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Krippendorff, K. (2016). Misunderstanding reliability. *Methodology*, *12*(4), 139–144. https://doi.org/10.1027/1614-2241/a000119

Krippendorff, K. (2019). Chapter 7: Recording/coding. In *Content analysis: An introduction to its methodology* (4th ed., pp. 128–156). Sage.

Kriss, L. A., Quick, B. L., Rains, S. A., & Barbati, J. L. (2022). Psychological reactance theory and COVID-19 vaccine mandates: The roles of threat magnitude and direction of threat. *Journal of Health Communication*, *27*(9), 654–663. https://doi.org/10.1080/10810730.2022.2148023

Krpan, D., & Dolan, P. (2022). You must stay at home! The impact of commands on behaviors during COVID-19. *Social Psychological and Personality Science*, *13*(1), 333–346. https://doi.org/10.1177/19485506211005582

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46–70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

LaVoie, N. R., Quick, B. L., Riles, J. M., & Lambert, N. J. (2017). Are graphic cigarette warning labels an effective message strategy? A test of psychological reactance theory and source appraisal. *Communication Research*, *44*(3), 416–436. https://doi.org/10.1177/0093650215609669

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, *46*(4), 352–367. https://doi.org/10.1037/0003-066X.46.4.352

Lee, S., & Lang, A. (2009). Discrete emotion and motivation: Relative activation in the appetitive and aversive motivational systems as a function of anger, sadness, fear, and joy during televised information campaigns. *Media Psychology*, *12*(2), 148–170. https://doi.org/10.1080/15213260902849927

Legros, S., & Cislaghi, B. (2020). Mapping the social-norms literature: An overview of reviews. *Perspectives on Psychological Science*, *15*(1), 62–80. https://doi.org/10.1177/1745691619866455

Lemmer, G., & Gollwitzer, M. (2017). The "true" indirect effect won't (always) stand up: When and why reverse mediation testing fails. *Journal of Experimental Social Psychology*, *69*, 144–149. https://doi.org/10.1016/j.jesp.2016.05.002

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H. M. Lefcourt (Ed.), *Research with the Locus of Control Construct* (pp. 15–63). Elsevier. https://doi.org/10.1016/B978-0-12-443201-7.50006-3

Lewin, K. (1939). Field theory and experiment in social psychology: Concepts and methods. *American Journal of Sociology*, *44*(6), 868–896. https://doi.org/10.1086/218177

Lieberman, M. D. (2010). Social cognitive neuroscience. In S. T. Fiske, D. T. Gilbert, & G.

  Lindzey (Eds.), *Handbook of social psychology* (pp. 143–193). John Wiley & Sons, Inc.

  https://doi.org/10.1002/9780470561119.socpsy001005

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain

  basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143.

  https://doi.org/10.1017/S0140525X11000446

Lu, F., & Sun, Y. (2022). COVID-19 vaccine hesitancy: The effects of combining direct and

  indirect online opinion cues on psychological reactance to health campaigns. *Computers in*

  Human Behavior, *127*, 107057. https://doi.org/10.1016/j.chb.2021.107057

McGuire, N. H., & Ball, H. (2022). Extending psychological reactance theory to include denial

  of threat and media sharing intentions as freedom restoration behavior. *Communication*

  Research Reports, *39*(3), 136–146. https://doi.org/10.1080/08824096.2022.2058480

Meltzer, C. E., Naab, T., & Daschmann, G. (2012). All student samples differ: On Participant

  selection in communication science. *Communication Methods and Measures*, *6*(4), 251–

  262. https://doi.org/10.1080/19312458.2012.732625

Miller, C. H., & Quick, B. L. (2010). Sensation seeking and psychological reactance as health

  risk predictors for an emerging adult population. *Health Communication*, *25*(3), 266–275.

  https://doi.org/10.1080/10410231003698945

Minich, M. (2023). Neural correlates of three putative mechanisms for communication-driven

  behavior change (Publication No. 30638291) [Doctoral dissertation, University of

  Wisconsin-Madison]. ProQuest Dissertations & Theses Global.

Minich, M., Chang, C.-T., Kriss, L. A., Tveleneva, A., & Cascio, C. N. (2023a). Gain/loss

  framing moderates the VMPFC's response to persuasive messages when behaviors have

personal outcomes. *Social Cognitive and Affective Neuroscience*, *18*(1). https://doi.org/10.1093/scan/nsad069

Minich, M., Chang, C.-T., Kriss, L. A., Tveleneva, A., & Cascio, C. N. (2023). Gain/loss framing moderates the VMPFC's response to persuasive messages when behaviors have personal outcomes. *Social Cognitive and Affective Neuroscience*, *18*(1), 1–9. https://doi.org/10.1093/scan/nsad069

Minich, M., Cotter, L., Kriss, L., Lu., L., Yang, S., & Cascio, C.N. (*Under review*). Pictorial warning labels reduce sharing intentions, alter self-relevance processes elicited by social media posts promoting cannabis edibles. Under review at *Journal of Communication*.

Minich, M., Tao, R., & Cascio, C. N. (2023b). Media and emotions: Insights from the neuroscience perspective. In J. G. Myrick & R. L. Nabi (Eds.), *Emotions in the digital world: Exploring affective experience and expression in online interactions* (pp. 76–93). Oxford University Press.

Mora, J.-D. (2016). Social context and advertising effectiveness: A dynamic study. *International Journal of Advertising*, *35*(2), 325–344. https://doi.org/10.1080/02650487.2015.1022975

Mutz, D. (2011). *Population-based survey experiments*. Princeton University Press.

Nabi, R. L. (2010). The case for emphasizing discrete emotions in communication research. *Communication Monographs*, *77*(2), 153–159. https://doi.org/10.1080/03637751003790444

National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, & Committee on the Health Effects of Marijuana: An Evidence Review and Research Agenda. (2017). The health effects of cannabis and cannabinoids: The current state of evidence and recommendations

for research. National Academies Press (US).

http://www.ncbi.nlm.nih.gov/books/NBK423845/

National Retail Federation. (2024). *Super Bowl Data Center*. https://nrf.com/super-bowl-data-center

Noar, S. M., Rohde, J. A., Barker, J. O., Hall, M. G., & Brewer, N. T. (2020). Pictorial cigarette pack warnings increase some risk appraisals but not risk beliefs: A meta-analysis. *Human Communication Research*, *46*(2–3), 250–272. https://doi.org/10.1093/hcr/hqz016

O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory*, *13*(3), 251–274. https://doi.org/10.1111/j.1468-2885.2003.tb00292.x

O'Keefe, D. J. (2015). Message generalizations that support evidence-based persuasive message design: Specifying the evidentiary requirements. *Health Communication*, *30*(2), 106–113. https://doi.org/10.1080/10410236.2014.974123

O'Keefe, D. J., & Hoeken, H. (2021). Message design choices don't make much difference to persuasiveness and can't be counted on—not even when moderating conditions are specified. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.664160

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, *56*(9), 717–734. https://doi.org/10.1037/0003-066X.56.9.717

Origlio, J., & Odar Stough, C. (2024). Locus of control and pre-pandemic depressive symptoms relate to psychological adjustment of college students to the COVID-19 pandemic. *Journal of American College Health*, *72*(2), 522–529. https://doi.org/10.1080/07448481.2022.2047699

Pennsylvania Department of Health. (2021, June 25). Department of health lifting universal masking order on June 28, urging Pennsylvanians to follow mask-wearing guidance where required [Press release]. https://www.media.pa.gov/Pages/Health-Details.aspx?newsid=1505

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Pew Research Center. (2016, April 26). A wider ideological gap between more and less educated adults. *Pew Research Center*. https://www.pewresearch.org/politics/2016/04/26/a-wider-ideological-gap-between-more-and-less-educated-adults/

Pittman, N. L., & Pittman, T. S. (1979). Effects of amount of helplessness training and internal–external locus of control on mood and performance. *Journal of Personality and Social Psychology*, *37*(1), 39–47. https://doi.org/10.1037/0022-3514.37.1.39

Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004

Poldrack, R. A. (2018). *The new mind readers: What neuroimaging can and cannot reveal about our thoughts* (Kindle). Princeton University Press.

Poole, M. S. (2013). On the study of process in communication research. *Annals of the International Communication Association*, *36*(1), 371–409. https://doi.org/10.1080/23808985.2013.11679140

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*(3), 390–395. https://doi.org/10.1037/0022-006X.51.3.390

Quick, B. L., Kam, J. A., Morgan, S. E., Montero Liberona, C. A., & Smith, R. A. (2015). Prospect theory, discrete emotions, and freedom threats: An extension of psychological reactance theory. *Journal of Communication*, *65*(1), 40–61. https://doi.org/10.1111/jcom.12134

Quick, B. L., Kriss, L. A., Morrow, E., Hartman, D., & Koester, B. (2024). A test of autonomy restoration postscripts to mitigate psychological reactance to an opt-out organ donor registry in the United States. *Health Communication*, *39*(9), 1714–1726. https://doi.org/10.1080/10410236.2023.2232607

Quick, B. L., Kriss, L. A., Rains, S. A., Sherlock-Jones, M., & Jang, M. (2023). An investigation into the portrayal of organ donation on Grey's Anatomy seasons 1 through 15. *Health Communication*, *39*(1), 195–204. https://doi.org/10.1080/10410236.2022.2163051

Quick, B. L., Shen, L. J., & Dillard, J. P. (2013). Reactance theory and persuasion. In J. P. Dillard & L. J. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 167–184). Sage. https://doi.org/http://dx.doi.org/10.4135/9781452218410.n11

Rains, S. A. (2013). The nature of psychological reactance revisited: A meta-analytic review. *Human Communication Research*, *39*(1), 47–73. https://doi.org/10.1111/j.1468-2958.2012.01443.x

Rains, S. A., & Richards, A. S. (2024). US state vaccine mandates did not influence COVID-19 vaccination rates but reduced uptake of COVID-19 boosters and flu vaccines compared to bans on vaccine restrictions. *Proceedings of the National Academy of Sciences*, *121*(8). https://doi.org/10.1073/pnas.2313610121

Rains, S. A., Colombo, P. M., Quick, B. L., & Kriss, L. A. (2022). State mask mandates and psychological reactance theory: The role of political partisanship and COVID-19 risk in mask adoption and resistance. *Social Science & Medicine*, *314*, 115479. https://doi.org/10.1016/j.socscimed.2022.115479

Ratcliff, C. L. (2021). Characterizing reactance in communication research: A review of conceptual and operational approaches. *Communication Research*, *48*(7), 1033–1058. https://doi.org/10.1177/0093650219872126

Real, K., & Rimal, R. N. (2007). Friends talk to friends about drinking: Exploring the role of peer communication in the theory of normative social behavior. *Health Communication*, *22*(2), 169–180. https://doi.org/10.1080/10410230701454254

Reynolds-Tylus, T. (2019). Psychological reactance and persuasive health communication: A review of the literature. *Frontiers in Communication*, *4*(56), 1–12. https://doi.org/10.3389/fcomm.2019.00056

Reynolds-Tylus, T., & Schill, D. (2023). An application of psychological reactance theory to college student voter registration and mobilization. In J. C. Baumgartner & B. I. Newman (Eds.), *Political marketing and the election of 2020*. Routledge.

Reynolds-Tylus, T., Bigsby, E., & Quick, B. L. (2020). A comparison of three approaches for measuring negative cognitions for psychological reactance. *Communication Methods and Measures*, 1–17. https://doi.org/10.1080/19312458.2020.1810647

Rhodes, N., Shulman, H. C., & McClaran, N. (2020). Changing norms: A meta-analytic integration of research on social norms appeals. *Human Communication Research*, *46*(2–3), 161–191. https://doi.org/10.1093/hcr/hqz023

Richard, Y., Tazi, N., Frydecka, D., Hamid, M. S., & Moustafa, A. A. (2023). A systematic review of neural, cognitive, and clinical studies of anger and aggression. *Current Psychology*, *42*(20), 17174–17186. https://doi.org/10.1007/s12144-022-03143-6

Richards, A. S., Bessarabova, E., Banas, J. A., & Larsen, M. (2021). Freedom-prompting reactance mitigation strategies function differently across levels of trait reactance. *Communication Quarterly*, *69*(3), 238–258. https://doi.org/10.1080/01463373.2021.1920443

Riddle, P. J., Newman-Norlund, R. D., Baer, J., & Thrasher, J. F. (2016). Neural response to pictorial health warning labels can predict smoking behavioral change. *Social Cognitive and Affective Neuroscience*, *11*(11), 1802–1811. https://doi.org/10.1093/scan/nsw087

Roditis, M. L., Jones, C., Dineva, A. P., & Alexander, T. N. (2019). Lessons on addiction messages From "The Real Cost" campaign. *American Journal of Preventive Medicine*, *56*(2). https://doi.org/10.1016/j.amepre.2018.07.043

Rosenberg, B. D., & Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, *4*(4), 281–300. https://doi.org/10.1037/mot0000091

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28. https://doi.org/10.1037/h0092976

Santiago, S., Mahoney, C., Murray, M. P., & Benoza, G. (2019a). "The Real Cost": Reaching at-risk youth in a fragmented media environment. *American Journal of Preventive Medicine*, *56*(2), S49–S56. https://doi.org/10.1016/j.amepre.2018.07.041

Santiago, S., Talbert, E. C., & Benoza, G. (2019b). Finding Pete and Nikki: Defining the target

   audience for "The Real Cost" campaign. *American Journal of Preventive Medicine*, *56*(2),

   S9–S15. https://doi.org/10.1016/j.amepre.2018.07.040

Seligman, M. E. (1974). Depression and learned helplessness. In R. J. Friedman & M. M. Katz

   (Eds.), *The psychology of depression: Contemporary theory and research*. Wiley.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. Freeman.

Shadel, W. G., Martino, S. C., Setodji, C. M., Dunbar, M., Scharf, D., & Creswell, K. G. (2019).

   Do graphic health warning labels on cigarette packages deter purchases at point-of-sale? An

   experiment with adult smokers. *Health Education Research*, *34*(3), 321–331.

   https://doi.org/10.1093/her/cyz011

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further

   exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6),

   1061–1086. https://doi.org/10.1037/0022-3514.52.6.1061

Shoenberger, H., Kim, E. (Anna), & Sun, Y. (2021). Advertising during COVID-19: Exploring

   perceived brand message authenticity and potential psychological reactance. *Journal of

   Advertising*, *50*(3), 253–261. https://doi.org/10.1080/00913367.2021.1927914

Shoham, V., Trost, S. E., & Rohrbaugh, M. J. (2004). From state to trait and back again:

   Reactance theory goes clinical. In R. A. Wright, J. Greenberg, & S. S. Brehm (Eds.),

   Motivational analyses of social behavior (pp. 167–186). Psychology Press.

   https://doi.org/https://doi.org/10.4324/9781410610089

Silvia, P. J. (2006). A skeptical look at dispositional reactance. *Personality and Individual

   Differences*, *40*(6), 1291–1297. https://doi.org/10.1016/j.paid.2005.11.013

Sittenthaler, S., Jonas, E., & Traut-Mattausch, E. (2016). Explaining self and vicarious reactance. *Personality and Social Psychology Bulletin*, *42*(4), 458–470. https://doi.org/10.1177/0146167216634055

Slater, D. M., Peter, J., & Valkenburg, P. M. (2015). Message variability and heterogeneity: A core challenge for communication research. *Annals of the International Communication Association*, *39*(1), 3–31. https://doi.org/10.1080/23808985.2015.11679170

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly*, *81*(1), 168–183. https://doi.org/10.1177/107769900408100112

Southwell, B. G. (2013). *Social networks and popular understanding of science and health: Sharing disparities* (Kindle). Johns Hopkins University Press.

Southwell, B. G., & Yzer, M. C. (2007). The roles of interpersonal communication in mass media campaigns. *Annals of the International Communication Association*, *31*(1), 420–462. https://doi.org/10.1080/23808985.2007.11679072

Southwell, B. G., & Yzer, M. C. (2009). When (and why) interpersonal talk matters for campaigns. *Communication Theory*, *19*(1), 1–8. https://doi.org/10.1111/j.1468-2885.2008.01329.x

Spellman, B. A., Gilbert, E. A., & Corker, K. S. (2018). Open science. In E.-J. Wagenmakers (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., volume 5, pp. 1–47). Wiley. https://doi.org/10.1002/9781119170174.epcn519

Sprengholz, P., Betsch, C., & Böhm, R. (2021). Reactance revisited: Consequences of mandatory and scarce vaccination in the case of COVID-19. *Applied Psychology: Health and Well-Being*, *13*(4), 986–995. https://doi.org/10.1111/aphw.12285

Sprengholz, P., Felgendreff, L., Böhm, R., & Betsch, C. (2022). Vaccination policy reactance: Predictors, consequences, and countermeasures. *Journal of Health Psychology*, *27*(6), 1394–1407. https://doi.org/10.1177/13591053211044535

Stuyt, E. (2018). High potency THC marijuana from the perspective of an addiction psychiatrist. *Missouri Medicine*, *115*(6), 482–487.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6312155/pdf/ms115_p0482.pdf

Sun, Y., & Lu, F. (2023). How misinformation and rebuttals in online comments affect people's intention to receive COVID-19 vaccines: The roles of psychological reactance and misperceptions. *Journalism & Mass Communication Quarterly*, *100*(1), 145–171.

https://doi.org/10.1177/10776990221084606

Super Bowl LVIII draws 123.7 million average viewers, largest TV audience on record. (2024, February 13). *Nielsen*. https://www.nielsen.com/news-center/2024/super-bowl-lviii-draws-123-7-million-average-viewers-largest-tv-audience-on-record/

Tal-Or, N. (2016). How co-viewing affects attitudes: The mediating roles of transportation and identification. *Media Psychology*, *19*(3), 381–405.

https://doi.org/10.1080/15213269.2015.1082918

Tops of 2019: Social TV. (2019, December). *Nielsen*.

https://www.nielsen.com/insights/2019/tops-of-2019-social-tv/

Tracy, S. J. (2013). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. Wiley Blackwell.

U.S. health warnings date pushed back again. (2022, November 15). *Tobacco Reporter*.

https://tobaccoreporter.com/2022/11/15/u-s-graphic-health-warnings-date-pushed-back-again/

Vezich, I. S., Katzman, P. L., Ames, D. L., Falk, E. B., & Lieberman, M. D. (2017). Modulating the neural bases of persuasion: why/how, gain/loss, and users/non-users. *Social Cognitive and Affective Neuroscience*, *12*(2), 283–297. https://doi.org/10.1093/scan/nsw113

Volkow, N. D., Swanson, J. M., Evins, A. E., DeLisi, L. E., Meier, M. H., Gonzalez, R., Bloomfield, M. A. P., Curran, H. V., & Baler, R. (2016). Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: A review. JAMA Psychiatry, 73(3), 292–297. https://doi.org/10.1001/jamapsychiatry.2015.3278

Wallston, K. A. (2005). The validity of the Multidimensional Health Locus of Control scales. *Journal of Health Psychology*, *10*(5), 623–631. https://doi.org/10.1177/1359105305055304

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the Multidimensional Health Locus of Control (MHLC) scales. *Health Education Monographs*, *6*(1), 160–170. https://doi.org/10.1177/109019817800600107

Weinhardt, C. B., & Ruckert, J. H. (2023). Internal locus of control predicts proenvironmental and COVID-19 health-related behaviors: A pilot study. *Ecopsychology*, *15*(4), 368–381. https://doi.org/10.1089/eco.2022.0016

Wicklund, R. A. (1974). *Freedom and reactance*. Lawrence Erlbaum.

Woodhouse, L. (2022, August 30). How weed became the new OxyContin. *Tablet Magazine*. https://www.tabletmag.com/sections/news/articles/how-weed-became-new-oxycontin-marijuana-psychosis-addiction

Wortman, C. B., & Brehm, J. W. (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. *Advances in Experimental Psychology*, *8*, 277–336. https://doi.org/10.1016/S0065-2601(08)60253-1

Xu, J. (2017). The impact of locus of control and controlling language on psychological

    reactance and ad effectiveness in health communication. *Health Communication*, *32*(12),

    1463–1471. https://doi.org/10.1080/10410236.2016.1230807

Yang, E. F., Kriss, L. A., & Sun, Y. (2023). Fun with frustration? TikTok influencers ' emotional

    expression predicts user engagement with COVID-19 vaccination messages. *Health*

    Communication, *00*(00), 1–16. https://doi.org/10.1080/10410236.2023.2259621

Yang, S., Cotter, L. M., Lu, L., Kriss, L. A., Minich, M., Liu, J., Silver, D., & Cascio, C. N.

    (2024). Countering online marketing and user endorsements with enhanced cannabis

    warning labels : An online experiment among at-risk youth and young adults. *Preventive*

    Medicine, *180*(December 2023), 107877. https://doi.org/10.1016/j.ypmed.2024.107877

Yang, S., Maloney, E. K., Tan, A. S. L., & Cappella, J. N. (2018). When visual cues activate

    moral foundations: Unintended effects of visual portrayals of vaping within electronic

    cigarette video advertisements. *Human Communication Research*, *44*(3), 223–246.

    https://doi.org/10.1093/hcr/hqy004

Youn, S., & Kim, S. (2019). Newsfeed native advertising on Facebook: Young millennials'

    knowledge, pet peeves, reactance and ad avoidance. *International Journal of Advertising*,

    38(5), 651–683. https://doi.org/10.1080/02650487.2019.1575109

Young, D. G., Rasheed, H., Bleakley, A., & Langbaum, J. B. (2022). The politics of mask-

    wearing: Political preferences, reactance, and conflict aversion during COVID. *Social*

    Science & Medicine, *298*, 114836. https://doi.org/10.1016/j.socscimed.2022.114836