Towards a functional action-based account of Moralization

Joshua Ju-suk Rhee

ORCID: 0000-0002-6245-7060

Doctor of Philosophy

December 2023

Faculty of Medicine, Dentistry and Health Sciences

Melbourne School of Psychological Sciences

Submitted in total fulfilment of the requirements of the degree Doctor of Philosophy.

Abstract

Morality has traditionally been considered to reflect a set of universal and objective values. However, more recent work investigating people's subjective moral beliefs have revealed that substantial and meaningful variance can exist within and between individuals in their tendency to see attitudes or actions as morally relevant. Moralization describes the process underlying within-person variance in attribution of moral significance to actions or attitudes. In this thesis we build on recent work investigating the mechanisms of subjective moralization by proposing the action-based moralization account – a framework that conceptualises moralization as functional response to action-based uncertainty in the context of threat-appraisal. To evaluate the utility of this account as a functional theory of moralization, we apply a multi-method approach to investigate the relationship between threat-appraisal and moralization. Specifically, we find evidence of a relationship between individual differences in threat-sensitivity and tendency for moralization. Through an ecological momentary assessment study, we also find that within-person variation in the experience threat-appraisal related affect predicts moralization. Finally, in two representative sample studies in the U.S. and U.K we find that individual differences in subjective threat-appraisal of COVID-19 is positively associated with moralization of COVID-19 preventative behaviours. Finally, we investigate the longitudinal relationship between moralization and uncertainty reduction in the context of a real-world political event. We find that moralization is associated with prediction confidence at time 1, and predicts subjective belief confirmation at time 2. Collectively, these findings provide initial support for the action-based moralization account, and highlight the value of functional theories as a framework for understanding the precursors and outcomes of moralization.

Declaration

I declare that:

(i)   the thesis comprises only my original work towards the Doctor of Philosophy

degree except where indicated in the preface;

(ii)   due acknowledgement has been made in the text to all other material used; and

(iii)   the thesis is fewer than 100 000 words in length, exclusive of tables, figures,

references and appendices.

1 December 2023, Joshua Ju-suk Rhee

Preface

For each study reported herein, experimental data was collected by the PhD candidate with assistance from members of the Ethics and Wellbeing Hub at The University of Melbourne. Six chapters of this thesis are presented in article format: Chapter 1 presents an article published by Social and Personality Psychology Compass on 5/12/2020; Chapter 2 presents an article in preparation for submission to Perspectives on Psychological Science; Chapter 3 presents an article submitted to Personality and Social Psychology Bulletin on 24/01/2022; Chapter 4 presents an article submitted to Proceedings of the Royal Society B on 06/01/2022; Chapter 5 presents an article in preparation for submission to Nature Human Behaviour; and Chapter 6 presents an article submitted to Journal of Social and Political Psychology on 23/05/2022. Each empirical manuscript was drafted by the PhD candidate, and was then revised and edited with the assistance of the manuscript's named coauthors. All co-authors have agreed to the use of these manuscripts in this thesis, and have provided signed copies of the co-author authorisation form. The work in this thesis was supported, an Australian Government Research Training Program Scholarship awarded to Joshua Rhee.

Publications, conference abstracts, and awards during candidature

Publications accepted during candidature forming part of this thesis

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Social and Personality Psychology Compass*, 13(12), e12511. https://doi.org/10.1111/spc3.12511

Publications submitted during candidature forming part of this thesis

Rhee, J.J., Tan, N.P., Ching, A. & Bastian, B (2023) *Toward an action-based moralization account.* Manuscript in Preparation.

Rhee, J.J., Tan, N.P., Ching, A. & Bastian, B (2022) *An Ecological Momentary Assessment of the Role of Incidental Affect on Moral Judgment in Daily Life.* Manuscript in Preparation.

Rhee, J. J., Donner, M. R., Klebl, C., McGrath, M. J., Song, J., Tan, P. J., Weaving, M., Van Reyn, C. & Bastian, B. (2022) *COVID-19 Threat-Appraisals are Positively Related to Moralization of COVID-19 Behaviours and Desire for Tight Cultural Norms.* [Manuscript Submitted for Publication]

Rhee, J. J. & Bastian, B. (2022) *Who is a Moralizer? Individual Differences in Sensitivity to Threat Predicts Greater Tendency for Moral Condemnation.* [Manuscript Submitted for Publication]

Rhee, J. J. & Bastian, B (2022) *Moralization Predicts Subjective Expectation Confirmation following release of Robert Mueller's report on his Investigations into Russia's involvement in the 2016 Presidential Election.* [Manuscript Submitted for Publication]

Conference Proceedings

Rhee, J.J., & Bastian, B. (2022) *An Ecological Momentary Assessment of the Role of Incidental Affect on Moral Judgment in Daily Life.* Abstract submitted to the 2022 annual conference of the Society for Personality and Social Psychology, Virtual.

Rhee, J.J., & Bastian, B. (2021). *Toward a Functional Threat-Based Account of Moralization: Insights from Individual and Cross-Cultural Differences, and Individuals' Experiences of Real-World Events.* Abstract submitted to the 2021 annual conference of the Society for Personality and Social Psychology, Virtual.

Rhee, J.J., & Bastian, B. (2020, July). *The Influence of Sensation Seeking and BIS on Moral Condemnation of Third-Party Actions.* Abstract submitted to the 19th annual conference of the European Association of Social Psychology, Krakow, Poland.

Rhee, J.J., & Bastian, B. (2019, April). *Who is a Moralizer? Individual differences in avoidance motivation and sensation seeking are associated with differences in the harshness of moral judgments.* Abstract submitted to the 48th annual conference of the Society for Australasian Social Psychology, Sydney, Australia.

Co-authored publications not forming part of this thesis

McHugh, C., Griffin, S. M., McGrath, M. J., **Rhee, J. J**., Maher, P. J., McCashin, D., & Roth, J. (2023). Moral Identity Predicts Adherence to COVID-19 Mitigation

Procedures Depending on Political Ideology: A Comparison Between the USA and New Zealand. *Political psychology*, 44(2), 337-360. https://doi.org/10.1111/pops.12838

Kothe, E. J., Ling, M., Mullan, B. A., **Rhee, J. J.**, & Klas, A. (2023). Increasing intention to reduce fossil fuel use: a protection motivation theory-based experimental study. *Climatic Change*, 176(3), 19. https://doi.org/10.1007/s10584-023-03489-1

Dakin, B., McGrath, M. J., **Rhee, J. J.** & Haslam, N. (2023) Broadened concepts of harm appear less serious. *Social Psychological & Personality Sciences.* https://doi.org/10.1177/19485506221076692

Dejonckheere, E., **Rhee, J.J.**, … Bastian, B. (*2023*) Perceiving societal pressure to be happy is linked to poor well-being, especially in happy nations. *Scientific Reports.*

Klebl, C., **Rhee, J. J.**, Greenaway, K. H., Luo, Y., & Bastian, B. (2021). Physical attractiveness biases judgments pertaining to the moral domain of purity. *Personality and Social Psychology Bulletin*, 01461672211064452. https://doi.org/10.1177/01461672211064452

Louis, W. R., Cibich, M., McGarty, C., Thomas, E.F., Amiot, C. E., Lizzio-Wilson, M., Weber, N., **Rhee, J. J.**, Davies, G., Rach, T., Goh, S., McMaster, Z., Muldoon, O., Howe, N., & Moghaddam, F. (in 2021) Failure Leads Protest Movements to Support More Radical Tactics. *Social Psychological and Personality Science.* https://doi.org/10.1177/194855062110372

Klebl, C., **Rhee, J.J.**, Greenaway, K., Luo, Y., & Bastian, B. (2020). Beauty goes down to the core: Attractiveness biases moral character attributions. *Journal of Nonverbal Behavior.* https://doi.org/10.1007/s10919-021-00388-w

Crone, D.L., **Rhee, J.J** & Laham, S.M. Developing brief versions of the Moral Foundations Vignettes using a genetic algorithm-based approach. *Behav Res* (2020). https://doi.org/10.3758/s13428-020-01489-y

Abeywickrama, R. S., **Rhee, J. J.**, Crone, D. L., & Laham, S. M. (2020). Why Moral Advocacy Leads to Polarization and Proselytization: The Role of Self-Persuasion. *Journal of Social and Political Psychology, 8*(2), 473-503. https://doi.org/10.5964/jspp.v8i2.1346

Haslam, N., Dakin, B. C., Fabiano, F., McGrath, M. J., **Rhee, J.**, Vylomova, E., Weaving, M., & Wheeler, M. A. (2020). Harm inflation: Making sense of concept creep. *European Review of Social Psychology*, *31*(1), 254-286. https://doi.org/10.1080/10463283.2020.1796080

Klebl, C., Greenaway, K. H., **Rhee, J. J.**, & Bastian, B. (2020). Ugliness Judgments Alert us to Cues of Pathogen Presence. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550620931655

Table of Contents

List of Figures

List of Tables

Overview
Towards a Functional Account of Moralization

Overview of the Thesis: Toward a Functional Account of Moralization

Throughout human history, people have evaluated and regulated the actions of themselves and others through the framework of moral "right" or "wrong" (Ayala, 2010; Ellemers et al., 2019; Ellemers & van den Bos, 2012; Tomasello & Vaish, 2013). Indeed, no lesser authority on evolutionary biology than Charles Darwin famously wrote: "of all the differences between man and the lower animals, the moral sense or conscience is by far the most important" (Darwin, 1872). Given this widespread view that morality is one of the defining aspects of the human condition, it should come as no surprise that psychological inquiry into morality is almost as old as the field of psychology itself (Dewey, 1902; Dewey & Tufts, 1932; May & Hartshorne, 1925). Such early psychological inquiries into morality started with an assumption about what constitutes moral content, and focused on how individuals arrive at the correct or incorrect moral answers or behaviour. However, more recent approaches have been directed toward observable patterns of cognition and behaviour that accompany anything that an individual subjectively considers moral (Lovett & Jordan, 2010; Skitka, 2010).

This more *subjective* approach to the study of morality has revealed two important insights about the nature and variability of people's moral beliefs. First, that variance can exist, not only in the types of issues that are seen as morally significant (Graham et al., 2009; Haidt & Graham, 2007), but also between individuals in the degree to which they moralize a given issue, object or action (Ryan, 2017; Skitka, 2010; Skitka et al., 2005). Second, that moral thinking represents a psychologically meaningful construct associated with a distinct set of behavioural and cognitive consequences (Skitka, 2005; Skitka, 2010; van Bavel et al., 2012). Collectively, these revelations have begun to highlight the "double-edged" nature of moralization. Within an environment of only like-minded believers, moralization can lead to increased willingness to sacrifice for the interests of the group (Haidt & Kesebir, 2010;

Swann et al., 2014), energization of collective action to advance group-goals (Agostini & van Zomeren, 2021; Turner-Zwinkels et al., 2017), and cooperation with group-members (Tomasello & Vaish, 2013). Yet, in an environment of belief heterogeneity – where substantial variance exists between individuals in their moralized beliefs – the very same mechanisms of moralization that facilitated ingroup cooperation, may engender out-group intolerance (Brandt et al., 2016), impede progress by decreasing willingness to compromise (Kovacheff et al., 2018; Ryan, 2017), and exacerbate conflict by increasing willingness to use violence (Mooijman et al., 2018).

Despite mounting evidence of (1) the potentially harmful consequences of moralization in certain contexts, and (2) the variability that can between individuals in the degree to which they moralize in said contexts, relatively little is known about the psychological mechanisms underlying the process of moralization itself (Skitka et al., 2021; Rhee et al., 2019; Skitka et al., 2018). The overarching objective of the current thesis is to contribute to understanding on this topic through the development of a new theoretical framework that looks at the cognitive and behavioural consequences of moralization through the perspective of how they may represent functional responses to an individual's circumstances and environment. To this end, the thesis is presented in two (2) theoretical chapters and four (4) empirical chapters:

Chapter 1 seeks to ground the contribution of the thesis in previous findings and theory by clarifying the taxonomy of the term *moralization* to reflect the way it has been used in psychological research and presenting a review of the current literature through this new taxonomy.

Chapter 2 presents a new theoretical framework for understanding the functions and motivations relevant to the process of moralization from the perspective of the individual.

This theoretical framework – which we call the action-based moralization account – forms the foundation of the empirical research program of the thesis.

Chapters 3 to 6 presents a series of seven (7) studies broadly directed toward establishing the initial feasibility of the action-based moralization account by empirically testing its core premises – namely, that individuals are likely to moralize in response to threat-appraisals; and that moralization functions to increase individuals' epistemic confidence. A more detailed overview of these empirical chapters is presented at the end of Chapter 2, where they are presented in the context of the action-based moralization account framework.

References

Agostini, M., & van Zomeren, M. (2021). Toward a comprehensive and potentially cross-cultural model of why people engage in collective action: A quantitative research synthesis of four motivations and structural constraints. *Psychological Bulletin*, *147*(7), 667–700. https://doi.org/10.1037/bul0000256

Ayala, F. J. (2010). The difference of being human: Morality. *Proceedings of the National Academy of Sciences*, *107*(Supplement 2), 9015–9022. https://doi.org/10.1073/pnas.0914616107

Brandt, M. J., Wetherell, G., & Crawford, J. T. (2016). Moralization and Intolerance of Ideological Outgroups. In *The Social Psychology of Morality*. Routledge.

Darwin, C. (1872). *The descent of man, and selection in relation to sex* (Vol. 1). D. Appleton.

Dewey, J. (1902). The evolutionary method as applied to morality. *The Philosophical Review, 11*(2), 107-124.

Dewey, J., & Tufts, J. H. (1932). Ethics (rev. ed.). *New York: H. Holt and company*.

Ellemers, N., & van den Bos, K. (2012). Morality in Groups: On the Social-Regulatory Functions of Right and Wrong. *Social & Personality Psychology Compass*, *6*(12), 878–889. https://doi.org/10.1111/spc3.12001

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The Psychology of Morality: A Review and Analysis of Empirical Studies Published From 1940 Through 2017. *Personality and Social Psychology Review*, *23*(4), 332–366. https://doi.org/10.1177/1088868318811759

Haidt, J., & Kesebir, S. (2010). Morality. In *Handbook of Social Psychology*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470561119.socpsy002022

Kovacheff, C., Schwartz, S., Inbar, Y., & Feinberg, M. (2018). The problem with morality: Impeding progress and increasing divides. *Social Issues and Policy Review*, *12*(1), 218–257.

May, M. A., & Hartshorne, H. (1925). Objective methods of measuring character. *The Pedagogical Seminary and Journal of Genetic Psychology*, *32*(1), 45-67.

Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature Human Behaviour*, *2*(6), 389–396.

Ryan, T. J. (2017). No Compromise: Political Consequences of Moralized Attitudes. *American Journal of Political Science*, *61*(2), 409–423.

Swann, W. B., Gómez, Á., Buhrmester, M. D., López-Rodríguez, L., Jiménez, J., & Vázquez, A. (2014). Contemplating the ultimate sacrifice: Identity fusion channels pro-group affect, cognition, and moral decision making. *Journal of Personality and Social Psychology*, *106*(5), 713–727. https://doi.org/10.1037/a0035809

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Turner-Zwinkels, F. M., van Zomeren, M., & Postmes, T. (2017). The moral dimension of politicized identity: Exploring identity content during the 2012 Presidential Elections in the USA. *British Journal of Social Psychology*, *56*(2), 416–436. https://doi.org/10.1111/bjso.12171

Chapter 1
Review of Current Definitions, Methods, and Evidence in Moralization Research

Chapter 1 Review of Current Definitions, Methods, and Evidence in Moralization Research

1.1.   Preface

This chapter presents a review and synthesis of the current landscape of moralization research. The manuscript contained in this chapter was published as an article in *Social and Personality Psychology Compass*.

The primary aim of this chapter was to review current approaches to the study and quantification of moralization, as well as influential theoretical perspectives on the topic. However, during the literature review process, we encountered significant challenges arising from fragmentation in the way that moralization was defined, measured, and theorised. Such fragmentation made it difficult to systematically identify relevant empirical findings and theoretical perspectives that could be captured under the umbrella of moralization. To overcome these challenges, we introduced a secondary aim – namely, to develop a taxonomic framework that reflected functional dimensions of variability present in the current literature on moralization.

Our final framework consisted of two taxonomic dimensions. The first dimension reflected how extant moralization research operationalized people's moral states – namely *via* judgments, attitudes and entities. The second dimension reflected the nature of the change in people's moral states that was considered to constitute moralization – specifically, (1) the psychological attachment of moral significance to an action, attitude or object (*Moral Recognition*); and (2) the process by which the moral significance of an action, attitude, or object becomes more extreme (*Moral Amplification*). Using this framework, we presented a review of the current landscape of moralization, including prevailing methodologies, empirical findings, and influential theoretical perspectives.

1.2.   Manuscript

The What, How and Why of Moralization: A Review of Current Definitions, Methods, and Evidence in Moralization Research

Joshua J. Rhee[1], Chelsea Schein[2], Brock Bastian[1]

[1]School of Psychological Sciences, University of Melbourne

[2]Department of Legal Studies and Business Ethics, Wharton School of Business, University of Pennsylvania

Author Note

Correspondence: Joshua J. Rhee, Melbourne School of Psychological Sciences, University of Melbourne, Victoria, 3010, Australia; rheej@student.unimelb.edu.au.

Word Count: 6547

Abstract

Moralization has major social and political implications. Although there is a depth of research on the nature and implications of moral attitudes and moral convictions, there has been less of a focus on the psychological processes by which actions, attitudes, or entities become moralized, or move from lesser to greater moral significance, and the research that does exist is highly fragmented. In the present paper we provide a two-factor structure for understanding the current state of research on moralization, categorizing extant moralization research by 1) whether it examines moral judgments, attitudes or entities, and 2) whether it captures moral recognition (the shift from neutral to moral), or moral amplification. Using this framework, we then consider the various routes through which moralization may occur, discuss emerging research on the influence that social norms can have on this process, and address future areas of research. Overall, we hope to provide some initial steps toward developing a more integrated framework for understanding moralization.

Key Words: moralization, morality, emotion, cognition, moral convergence, moral amplification, norms

The What, How and Why of Moralization: A Review of Current Definitions,

Methods, and Evidence in Moralization Research

Morality changes over time. This change can occur at the micro/individual level, such as when someone embraces ethical vegetarianism or changes their attitudes towards abortion or having a baby out of wedlock. It can also occur at the macro, societal level, such as the marked shift in moral attitudes toward gay marriage over the last 15 years (Gallup, 2019). Although moral change certainly occurs with potentially major social and interpersonal consequences, the academic study of moralization, or increases in the *degree* to which moral relevance is attached to issues, actions or entities, is still relatively new and fragmented.

Decades of research have documented the nature and implications of categorizing an act as morally relevant (Rozin, 1999; Ryan, 2014; Skitka, 2010; Skitka, Bauman, & Sargis, 2005; van Bavel, Packer, Haas, & Cunningham, 2012). Acts that are viewed as morally relevant are perceived by individuals as universally wrong or right (Skitka et al., 2005; van Bavel, Packer, Haas, & Cunningham, 2012), akin to objective facts (Goodwin & Darley, 2008) and beyond compromise (Skitka, 2010). Once an act is labelled as immoral, this categorization has important social implications. Moral conviction regarding a collective cause is a core motivator for individuals to engage in collective action (van Zomeren, 2013; van Zomeren, Postmes, Spears, & Bettache, 2011) and leads to increased hostility toward ideological outgroups (Brandt, Crawford, & Van Tongeren, 2017; Brandt et al., 2016).

Despite the fact that moral beliefs are often perceived as akin to objective facts (Goodwin & Darley, 2008; Goodwin & Darley, 2012), considerable variability can exist between societies in the types of issues that are considered morally relevant (e.g. Graham, Haidt, & Nosek, 2009; Vauclair & Fischer, 2011). Moral theories have explained cultural differences by appealing to variation in Moral Foundations (Graham et al., 2009; Haidt & Graham, 2007), relationship models (Rai & Fiske, 2011), or behavioural inhibition tendencies

(Janoff-Bulman, Sheikh, & Hepp, 2009). These theories each provide comprehensive approaches to modelling moral diversity. However, these models have not empirically demonstrated how attitudes change over time.

As the number of studies on moralization have increased (e.g. Brandt et al., 2016; Feinberg, Kovacheff, Teper, & Inbar, 2019; Skitka, Wisneski, & Brandt, 2017), it has also become increasingly apparent that the literature contains fragmentation in the way that moralization is defined, conceptualised, and measured. This paper aims to provide a general review of the current landscape of moralization research. More specifically, we review current methods applied to the study of moralization and discuss how the phenomenon has been studied in terms of judgments, attitudes and entities (see *Figure 1*). We also discuss inconsistencies in how moralization has been defined across these different foci and observe that – while not typically acknowledged – moralization has been studied in two forms: moral recognition and moral amplification. Finally, applying this conceptual distinction between moral recognition and moral amplification (see *Figure 1*), we conduct a review of the empirical literature to date and discuss the evidence for psychological mechanisms pertaining to these two forms of moralization. Our aim is not to advocate for any particular approach to studying moralization, but to provide a broadened platform from which to understand and study the general processes underlying moralization. To this end, we also propose a broader definition of moralization which better encapsulates the various ways in which it has been researched and understood – that is, an increase in the *degree* to which moral relevance is attached to issues, actions or entities.

Past and Present Definitions of Moralization

The term moralization appeared intermittently in a number of social psychology publications in the early 20th century, with perhaps the most notable examples coming from the early efforts of John Dewey to define a scientific framework for the study of morality

(Dewey, 1902; Dewey & Tufts, 1932). A key part of Dewey's framework was the idea that data on how norms and moral goals were formed in the past could shed light on the processes of moralization (Dewey, 1902). The term moralization was also used more formally as a research construct within the framework of Kohlberg's cognitive-development account, in which it was used to describe an individual's progress through discrete stages of moral reasoning ability (Kohlberg, 1976).

By far the most commonly used definition of moralization in research today (Brandt et al., 2016; Lovett & Jordan, 2010; Skitka, Wisneski, & Brandt, 2017) comes from the seminal work of Rozin and colleagues, who investigated changes in moral attitudes toward smoking and vegetarianism (Rozin, 1997, 1999; Rozin, Markwith, & Stoess, 1997; Rozin & Singh, 1999). Based on observations from his own research, Rozin (1997) noted that the moral status of an activity may ebb and flow over time, and that this constituted a meaningful psychological phenomenon that merited scholarly investigation and analysis. Within this conceptualization, Rozin defined moralization as a process which "*involves the acquisition of moral qualities by objects or activities that previously were morally neutral*" (Rozin, 1999; Rozin, Markwith, & Stoess, 1997; Rozin & Singh, 1999). There are two important points to note in this definition. First, Rozin explicitly stipulates that moralization starts with a morally neutral state. Next, Rozin's approach to moralization, is both descriptive and value-neutral (Lovett & Jordan, 2010), aiming to understand the psychological precedents or antecedents of an individual attributing subjective moral significance to something previously considered non-moral.

As the field of moral psychology has blossomed, researchers have drawn on Rozin's definition of moralization as a means of studying moral change (Brandt et al., 2016; Lovett & Jordan, 2010). However, developments in the field of moral psychology research – in particular, an increased focus on how moral values vary at the level of the individual (Bloom,

13

2010; Haidt, 2007) – have led to the use of a variety of methods and approaches for conceptualising and measuring individuals' moral positions, in turn leading to some ambiguity in how moralization should be understood.

In the next section, we present an overview of some of the notable methods that have been applied to the study of moralization, before engaging in a more formal discussion of the definitional ambiguities that have arisen and how to navigate them.

Methods of Studying Moralization

Empirical work on moralization has tended to apply one of three broad approaches to the study of moral change: cross-sectional; experimental; and longitudinal. Cross-sectional methods were used extensively in early studies of moralization by Rozin and colleagues (Rozin et al., 1997; Rozin & Singh, 1999). For example, Rozin and Singh (1999), captured the process of moralization by sampling three different generations of participants (students, their parents, and their grandparents), and having them rate their current attitudes regarding cigarette smoking, such as whether "cigarette smoking is immoral," or "cigarette smoking is disgusting" then asking them how they would have responded to the attitude items 20 or 40 years prior (Rozin & Singh, 1999).

The most common contemporary method of studying moralization has been to apply experimental manipulations which test specific mechanisms predicted to influence moral processes. These studies infer that factors which are able to produce significant group differences in the moral wrongness or perceived moral significance of an action or issue may also bring about changes in moral values within the individual (Horberg, Oveis, Keltner, & Cohen, 2009; Wisneski & Skitka, 2016). However, because moralization research is interested in changes in individuals' moral perceptions, the highest standard of evidence for moralization is considered to come from longitudinal designs, in which measures of individuals' current moral position, and relevant predictors of interest are measured at

multiple timepoints (Brandt et al., 2016; Brandt, Wisneski, & Skitka, 2015; Feinberg et al., 2019; Skitka et al., 2017).

Cutting across these methodological differences, studies of moralization have also differed in the way that they conceptualised and how they have measured people's moral states. Here we overview three notable approaches to measuring individuals' moral states that currently appear in the literature – namely, via judgments, attitudes and entities (see *Figure 1*).

Figure 1

Illustrations of moral recognition and moral amplification across judgments, attitudes and entities. Attitudes item is taken from: Skitka, L. J., Bauman, & Sargis, E. G. (2005). Entities item is taken from: Crimston, Bain, Hornsey, & Bastian (2016).
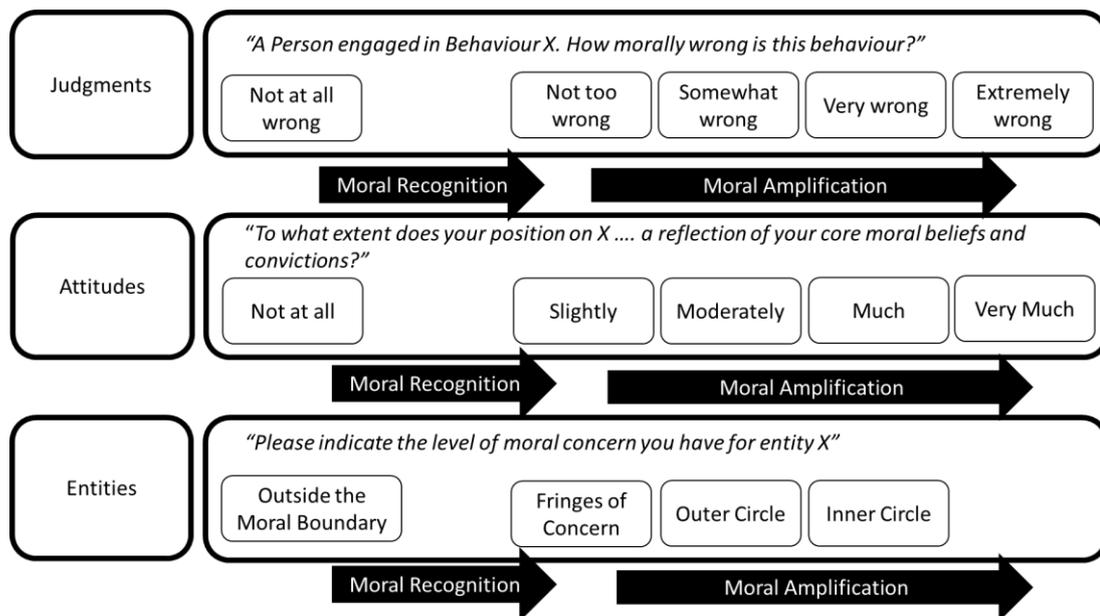
Moral Judgment of Specific Actions

The most common approach to observing individuals' moral states has been to measure their moral judgments of specific actions. This approach generally starts by presenting participants with an example of hypothetical (e.g. Clifford, Iyengar, Cabeza & Sinnott-Armstrong, 2015) or real-world (e.g. Lindstrom, Jagard, Selbing & Olsson, 2017) behaviour that could be construed as a moral violation (e.g. "*Person X takes money out of a wallet they find on the street*"), and then prompting them to indicate how morally wrong they

perceive the action to be (usually on a scale from "*not at all morally wrong*" to *"extremely morally wrong"*). Moral judgment approaches can be particularly well-suited to investigating how reasoning ability (Royzman, Landy, & Goodwin, 2014), heuristics (Haidt, 2001; Haidt, Koller, & Dias, 1993), social-learning (Davidson, Turiel, & Black, 1983), and transient contextual factors such as affective states (Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005) may influence individuals' engagement with moral scenarios. However, one of the limitations in operationalizing moralization by differences in the severity of moral condemnation of a single action is that it provides little information about what exactly has been *moralized* in a more general sense (i.e. beyond condemning the individual action), making it difficult to determine how this change might influence future judgments or behaviour (Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015).

Moral Attitudes

An increasingly prominent method of studying moralization is to investigate changes in the degree to which individuals ascribe moral significance to their attitudes on issues (e.g. gun control) or behaviours (e.g. smoking; Skitka, 2010; Skitka et al., 2005; Skitka et al., 2017). While moral judgments ask participants to indicate the moral wrongness of a specific instance of behaviour, moral attitude approaches focus on the application of morality at a more abstract level (for illustration of differences see *Figure 1*). For example, while a judgment measure may ask participants to indicate how morally wrong they perceive a person smoking to be, an attitude measure would ask how morally wrong they consider smoking as a general behaviour (Rozin & Singh, 1999).

The most extensive recent research into the moralization of attitudes has come via an extension of the moral conviction program (Skitka, 2010; Skitka et al., 2005; Wisneski & Skitka, 2016), where there has been a focus on factors that may account for upward shifts in participants' self-reported level of moral conviction regarding their position on an attitude

16

object (e.g. one's position on abortion). The defining characteristic of moral conviction is how central the attitude is to individuals' core moral beliefs and convictions, or fundamental beliefs about right and wrong (Skitka et al., 2005).

This approach to measuring moral attitudes can be particularly advantageous in identifying factors which contribute to moralization in the context of highly polarized issues such as abortion (Wisneski & Skitka, 2017) or political elections (Brandt, et al., 2016), allowing the same scale to be used to measure moral changes occurring for both opponents and supporters (e.g. "to what extent does your position on abortion reflect your core moral beliefs and convictions?").

Entities

An emerging perspective in the study of moralization has been to investigate differences in the level of moral concern for *entities* (Crimston, Bain, Hornsey, & Bastian, 2016; Crimston, Hornsey, Bain, & Bastian, 2018). This conceptualisation of moralization is defined by a focus on whether an entire category of groups (e.g. Criminals, Outgroups, Family Members) or entities (e.g. Redwood Trees, Dolphins, Rivers) are perceived as generally deserving of moral concern (Crimston et al., 2016) or can be considered moral patients/victims (Schein & Gray, 2014, 2016). The advantage of this approach is that it can potentially simultaneously predict the attachment of, or increase in, moral significance of a number of attitudes or actions (for example items see *Figure 1*). For example, once it was recognized that animals were capable of suffering and classified as potential moral patients, a number of attitude objects such as animal product testing, animal consumption, and animal mental health also came to be seen as moral issues (Rozin et al., 1997; Schein & Gray, 2016).

<center>**Revisiting Rozin's Definition**</center>

Perhaps because the various methods outlined in the previous section are also attached to different theoretical positions (for example, much of the research on moralization in terms

of changes in moral judgment emerged as an extension of the social intuitionist model; Haidt, 2001), there has been little formal engagement with how these methods are related to one another, and in turn, what empirical insights arising from each method may mean for understanding moralization as a general process. As a consequence, there remains a certain degree of inconsistency in the type of phenomena that are viewed as constituting moralization.

While most contemporary researchers cite Rozin's definition of moralization (Feinberg et al., 2019; Horberg, Oveis, & Keltner, 2011; Skitka et al., 2017) – i.e. the acquisition of moral qualities by objects or activities that previously were morally neutral – there is a great deal of variability in the type of empirical evidence seen to constitute this phenomenon. Most notably, researchers investigating changes in moral judgment have tended to make a distinction between *moralization* and *moral amplification*. According to this distinction, something would be considered moralization if a morally neutral act enters the moral sphere, while a shift from seeing a slightly wrong act as more wrong would be considered moral amplification (Avramova & Inbar, 2013; Landy & Goodwin, 2015; Pizarro et al., 2011).

On the other hand, recent research on moralization using the moral conviction paradigm has tended to employ a less restrictive standard when using the term moralization. For example, a number of key studies investigating moralization using the moral conviction approach have reported increases from "moderate" levels of moral conviction regarding an issue (Brandt, Wisneski, & Skitka, 2015; Wisneski & Skitka, 2016), making it difficult to suggest that the study demonstrated a change from the non-moral to moral domain. Even among studies of moral judgment, the use of the term moralization has tended to be inconsistent. For example, while citing Rozin (1997), Horberg et al. (2009) described

moralization as a process '*amplify[ing]* the importance of different moral domains during moral judgments', as opposed to judging a previously non-moral action as morally wrong.

Such differences in the use of the term moralization make it difficult to see how findings that have applied these different approaches may be reconciled. Should research that fails to demonstrate a change from non-moral to moral be completely excluded from any general account of moralization? We believe that to do so would unduly restrict the range of empirical work that could be considered in the formation of more general theories of moralization. Yet, we also believe that the lack of conceptual clarity that is currently present in moralization research needs to be addressed.

Thus, for the purposes of the review below, we use a broader definition of moralization– that is, increases in the *degree* to which moral relevance is attached to issues, actions or entities. We believe this broadening of the definition of moralization will mean that a greater range of empirical findings can be drawn on to leverage insight into the processes of moralization. However, for the sake of conceptual and theoretical clarity, and because it may provide greater insight into the processes of moralization, we also refer to two kinds of phenomena within this broad concept of moralization: (1) the psychological attachment of moral significance to an action, attitude or object (*Moral Recognition*); and (2) the process by which the moral significance of an action, attitude, or object becomes more extreme (*Moral Amplification*).

Evidence for Moral Recognition: A Process of Attachment and Detachment of Moral Significance

As we note, we use the term *moral recognition* to describe the psychological process of attaching moral significance to a given action, attitude or entity. Although the attachment (or detachment) of moral significance to a particular issue has often been considered in terms of long-term socio-cultural change (e.g. societal views on smoking; Rozin, 1999; Schein &

19

Gray, 2018), there is now increasing evidence to suggest that it may be a more dynamic process. Notably, at the level of making moral judgments, there is now a growing field investigating the phenomenon of *ethical fading* (Tenbrunsel & Messick, 2004) – where the ethical dimensions of certain decisions are not recognized (Sezer, Gino, & Bazerman, 2015). Such studies have revealed that situational and cognitive factors such as the salience of monetary interests (Gino & Mogilner, 2014) and previous engagement in the same behaviour (Shu & Gino, 2012), can lead to moral detachment, whereby people no-longer see the moral significance of a given action. This work indicates that moral recognition may be influenced by contextual factors as well as individual motivations, and that the attachment and detachment of moral significance may fluctuate over time and across situations within individuals. In concrete terms, for the purposes of this review, we have included studies under the umbrella of moral recognition if they contain evidence of changes in moral position *from* a complete lack of moral significance (see *Figure 1*).

Moral recognition through incidental emotions. As previously mentioned, only a handful of empirical studies on moralization have demonstrated the process of moral recognition (Avramova & Inbar, 2013). The most promising early work came from studies that were heavily influenced by the social intuitionist model of morality (Haidt, 2001), which posits that intuitions and emotions (rather than deliberative reasoning) are the preliminary drivers of moral judgment. This account was tested using "incidental emotion" paradigms (Bodenhausen, 1993), where participants were usually asked to make a moral judgment after being induced to feel an emotion (most often disgust) by a completely irrelevant stimulus (e.g. fart gas, pictures of dirty rooms).

For example, Wheatley and Haidt (2005) hypnotically primed one group of participants to feel disgust after reading the word "often", and the other group to feel disgust after reading the word "take", and found that participants were harsher in their moral

condemnation of actions portrayed in vignettes containing their specific disgust-priming word. Although most of the vignettes presented to participants by Wheatley & Haidt (2005) portrayed moral violations, this study is most widely cited for demonstrating moral recognition by showing increased moral condemnation for the following non-moral vignette:

''Dan is a student council representative at his school. This semester he is in charge of scheduling discussions about academic issues. He [tries to take/often picks] topics that appeal to both professors and students in order to stimulate discussion.''

More specifically, when no disgust-inducing words were present, participants rated this vignette $M = 2.7$, on a 100-point scale (0 = *'not at all morally wrong'* and 100 = *'extremely morally wrong'*), however when disgust inducing words were present moral condemnation rose to $M = 14.0$.

Despite the initial promise of incidental affect accounts, there has been very little subsequent work clearly demonstrating moral recognition following an incidental affect induction. In a recent meta-analysis conducted by Landy and Goodwin (2015), it was found that incidental affect inductions did appear to have a small, but significant effect on moral recognition ($d = .21$, $p = .01$, 95% CI [.05, .37]). However, the interpretation of this effect size was limited by the small number of studies ($k = 13$). Further, of these studies, seven were unpublished, and the weighted mean effect for these unpublished studies was non-significant ($d = .14$, $p = .18$, 95% CI [-.06, .34]), suggesting some presence of publication bias. Although such findings do not unequivocally discount incidental affect as a mechanism leading to moral recognition, they do make the importance of this mechanism unclear.

Moral recognition through moral emotions and moral piggybacking. Feinberg, Kovacheff, Teper, & Inbar (2019) recently conducted the only longitudinal study to date which demonstrated a process that could be viewed as *moral recognition*. Of particular note was Study 3, where participants were presented with an initial screening question: "Overall,

how much do you believe eating meat is immoral?" and were only included in the study if they indicated a "1" or "*not at all*", thus ensuring that only the phenomena of moral recognition was observed. These participants then took part in a seven-session study, where each session was conducted once every four to five days. In alternating sessions, participants were either: measured on their current moral attitudes regarding meat eating, and key hypothesised predictors of moralization; or presented with videos that viscerally portrayed the suffering of animals inflicted by the meat-consumption industry, which was intended as a way of inducing moralization.

Cross-lagged structural equation modelling analyses revealed two important predictors of moralization at subsequent timepoints: moral-emotions and moral-piggybacking. More specifically, controlling for participants' already existing level of moralization: the degree to which participants felt moral emotions (e.g. disgust) at the thought of eating meat at a given time point; and moral-piggybacking (e.g. the extent to which killing animals for meat was associated with larger moral principles they live by) positively predicted increases in the moral significance they attributed to meat-eating attitudes. It was also found that moral-emotions and moral-piggybacking fully mediated the effect of other predictors of moralization such as the perceived suffering inflicted on animals by the meat-industry, and the perception that animals and humans are similar.

Summary. The small amount of empirical work directly examining the process of moral recognition (as opposed to amplification) means that there is still little that is known about the underlying processes. While the evidence suggests the incidental affect may play a smaller role than initially thought (Landy & Goodwin, 2015), the work of Feinberg et al. (2019) identifies two promising mechanisms for moral recognition research – namely moral-emotions, and moral-piggybacking.

Evidence for Moral Amplification: Changes in Extremity

While the term 'moral amplification' is well-established within studies of moral judgment (Landy & Goodwin, 2015), as we note, it has not been formally considered as a process relevant to the study of moralization. By linking it to work on moralization, we offer a broader perspective on what kinds of targets moral amplification may be applied to. We define moral amplification as changes in the degree of extremity or importance of an already moral position. In more concrete terms, for the purposes of this review, we have included studies under the umbrella of moral amplification if the baseline measure of participants' moral position was above the lowest point in the relevant scale (see *Figure 1*).

Amplification through incidental emotions. In contrast to studies of moral recognition, there has been a number of studies demonstrating that incidental affect – in particular disgust – can lead to moral amplification of wrongness judgments (Johnson et al., 2016; Pizarro et al., 2011; Schnall, Haidt, Clore, & Jordan, 2008). However, a recent meta-analysis by Landy and Goodwin (2015) found these effects to be small ($d = .11$, $p = .002$, 95%CI [.04, .19]). Interestingly, it was also found that studies inducing olfactory (i.e. bad odors) or gustatory (i.e. bad tastes) forms of disgust appeared to demonstrate the most reliable moral amplification effects ($d = .37$, 95% CI [.15, .60]), compared to other forms of disgust induction such as visual (e.g. a video showing a disgusting toilet), or imaginary (e.g. recalling an event that made them feel disgusted).

The strong effect of olfactory or gustatory disgust inductions appears consistent with pathogen-avoidance accounts, which predict that the sensory modalities of smell and taste may be strongly associated with disgust (Schaller & Park, 2011). Although, it should also be noted that the number of olfactory and gustatory disgust inductions considered in the meta-analysis were much smaller ($k = 8$) than visual ($k = 35$) or imaginary ($k = 17$) inductions. Further, given the small size of these effects, even if incidental affect does influence moral wrongness judgments, it is unlikely that it is a primary driver of moralization.

Amplification through attitude-specific emotions. Notwithstanding the limited effect of incidental affect inductions, there remains strong empirical support for the general relevance of emotions in moral amplification where the emotion can be consciously tied to a specific attitude object. This was demonstrated in a longitudinal study conducted by Brandt et al. (2015), who sought to investigate whether certain emotions or beliefs could predict changes in the degree of moral conviction that participants felt for their preferred and non-preferred presidential candidates in the weeks leading up to the 2012 U.S. Presidential election. Contrary to influential accounts of moralization (e.g. Schein & Gray, 2018), it was found that beliefs regarding the potential harms (for non-preferred) or benefits (for preferred) of a given candidate at Time 1 was not a significant predictor of changes in moral conviction for those candidates. Instead, it was found that changes in moral conviction regarding candidates between Time 1 and Time 2 were predicted by enthusiasm (indicated by emotions such as happy or excited) at Time 1 (for preferred candidates), and hostility (indicated by emotions such as anger or disgust) at Time 1 (for non-preferred candidates).

Wisneski and Skitka (2016) sought to experimentally investigate the potential influence of incidental affect by presenting participants with disgust inducing images that were: abortion relevant (e.g. images of foetuses); related to a moral issue but not to abortion (e.g. of suffering animals); or non-moral (e.g. overflowing toilets). The mean level of moral conviction in each of the image groups were then compared with a control condition where participants simply viewed images of mundane objects (e.g. tables, pencils). It was found that participants in both the non-abortion related disgust conditions did not report significantly higher levels of moral conviction regarding abortion compared to participants in the control condition. On the other hand, participants in the abortion-related disgust condition did report elevated levels of moral conviction regarding abortion compared to the control group. Furthermore, it was found that this difference in moral conviction was fully mediated by

24

differences in the self-reported level of disgust experienced by participants, but not by other commonly proposed mechanisms of moralization such as differences in self-reported anger, or the perceived harmfulness of abortion (Skitka et al., 2017). Such findings suggest that it was the specific evocation of disgust in the context of attitude-relevant images that led to moral amplification.

Amplification through moral communication. Another promising area of research on moral amplification has examined the influence of receiving messages framed using moral language. For example, in a recent study, Clifford (2018) found that presenting participants with a persuasive message argued in terms of both disgust (e.g. that farmed animals live around their own urine and faeces) and harm (e.g. that pig-farmers break the tips of pigs' teeth to prevent biting) resulted in increased levels of moral conviction compared to a control group (where participants were not presented with any messages). Importantly, it was found that these differences in moral conviction persisted at a two-week follow up. Further, participants who had been exposed to the harm-based persuasion frame – but not the disgust-based frame – reported significantly higher levels of support for food-related policies and said that they would be more upset if a family member disagreed with them on the issue of natural-food, when compared to a control condition.

A similar process was demonstrated in a study of children by Rottman, Young, and Kelemen (2017). In this study, participants were presented with a series of 12 pictures portraying aliens engaging in novel behaviours (e.g. sticking cotton balls in their ear). Depending on their assigned condition, participants were also induced to experience disgust via exposure to fart spray; presented with a description which framed the behaviour as disgusting (e.g. "Acting like this is really gross"); or presented with a description framing the behaviour as anger-inducing (e.g. "Acting like this is really irritating."). It was found that, compared to participants in a control condition (who saw only the images), participants who

were presented with descriptions of the behaviours in "disgusting" or "anger-inducing" terms were more likely to rate the behaviours as "wrong", however, this difference was not observed for participants who were induced to feel disgust using fart-spray.

Both of these studies demonstrate that receiving communication from others which makes salient the moral features of a given attitude or action may be an important avenue for moral amplification. Indeed, at least when it comes to children, it appears that receiving communication that expresses others' moral emotions (e.g. disgust or anger) regarding an action may be more strongly linked to moral amplification than the general induction of disgust.

Amplification through descriptive norms. Social factors may play a role in moralization beyond communication processes. Group norms refer to social rules or standards for behaviour (Terry & Hogg, 1996), which can either be descriptive or injunctive in nature. Descriptive norms refer to perceptions of behaviours that are performed by the majority of others within a relevant referent group, while injunctive norms are perceptions of behaviours that the majority of others approve or disapprove of (Cialdini et al., 1991). Although moralization is often conceptualised as the conversion of descriptive regularities in behaviour to injunctive norms (Morris & Liu, 2015), recent research has found that information regarding behaviours that constitute the descriptive norm may have a powerful effect on the degree to which they are seen as moral. For example, within a public goods game paradigm (a multi-player game where players that are initially endowed with a certain sum of money can choose to invest some of their money to a shared fund) Lindström et al. (2017) manipulated the perceived "commonness" of sharing behaviour by varying the rate at which fictional players were represented as contributing to the shared fund across multiple rounds. It was found that, in conditions where sharing behaviour was common (i.e. players shared 80% of the time), participants were more likely to condemn the failure to share and

were more likely to incur personal cost to punish the behaviour (see also: Monroe, Dillon, Guglielmo, & Baumeister, 2018).

Summary. Incidental emotions appear to have small effects on moral amplification of judgments, although in the case of disgust, this may be moderated by the way in which it is induced, with some evidence suggesting that olfactory or gustatory inductions of disgust may have a particularly strong influence on moral amplification. Although incidental emotions appear to have no effect on attitudes, moral amplification does appear to be predicted by the activation of emotions that can be tied to the relevant attitude. Converging evidence also suggests that framing behaviours or issues in emotional or harm-based terms can lead to moral amplification for both judgments and actions. Finally, an emerging body of empirical evidence suggests that moral amplification may also be driven by perceptions that a given action represents a descriptive norm – that is, behaviours that are performed by the majority of individuals.

Future Directions

In this review of the growing field of moralization research we have sought to contribute to future research and theorizing by bringing together the extant literature within an organizing conceptual framework. Here we lay out three specific directions for future research that we believe are highlighted by our review.

Interplay between Moral Judgments, Attitudes and Entities

Our discussion of moralization across three different domains – namely, judgments, attitudes and entities may provide a platform for future research seeking to identify mechanisms that are most applicable to a specific domain of moralization. In our review of the empirical evidence on moralization across these domains, it is apparent that while certain mechanisms may apply to all three (e.g. moral emotions, and morally-framed communication), others may be domain-specific. For example, as discussed above, the

inducement of incidental disgust appears to have a small but significant effect on the moral amplification of judgments but have no influence on the moral amplification of attitudes.

Greater understanding of which mechanisms are most relevant to moralization of a given domain; and how moralization occurring in a specific domain may influence other relevant domains, may lead to particularly important insights for political and social movements seeking to bring about widespread changes in moral concern. For example, where there is difficulty shifting individuals' level of moral concern for out-groups, the path of least resistance may be to start by targeting relevant moral judgments, or *vice versa*.

Moral Amplification versus Moral Recognition

The introduction of a formal conceptual framework for separating moral recognition from moral amplification, and thereby the inclusion of existing research on moral amplification within the conceptual domain of moralization, may also lead to new directions for moralization research. To be clear, by introducing this framework, we do not make the claim that the two phenomena are empirically distinct. Indeed, it is highly likely that the psychological factors influencing moral amplification may also be relevant for moral recognition – as our review suggests. However, conceptualizing recognition and amplification as distinct processes may lead to future research into factors that are more strongly applicable to one process or the other. For example, it is possible that cognitive mechanisms underlying moralization such as moral-piggybacking (Feinberg et al., 2019), or appraisals of harm (Schein & Gray, 2018) may be more relevant to the process of moral recognition (see also, Skitka et al., 2017), while emotional mechanisms may play a greater role in the process of moral amplification.

Integrating Insights from De-moralization

Finally, a generative direction for future research and theorizing may be to integrate insights arising from research on processes underlying de-moralization. Indeed, Rozin

originally defined the concept of moralization by reference to its inverse, which he termed *amoralization* (Rozin, 1997; Rozin et al., 1997), suggesting that he envisaged one process as informed by the other. In this regard, one particularly informative area of demoralization research may be the growing field investigating the phenomenon of *ethical fading* (Tenbrunsel & Messick, 2004), which documents the various factors which may lead people to no-longer see the moral significiance of a given action. (Gino & Mogilner, 2014; Sezer, Gino, & Bazerman, 2015; Shu & Gino, 2012). Such research may inform the future study of moralization by identifying psychological barriers that may prevent moralization from occurring, even when moralization-inducing factors are present. This approach was recently tested by Feinberg et al. (2019) who proposed that an individual's position on a moral attitude may represent an equilibrium between processes "pushing" toward greater moralization (e.g. moral emotions), and "pull" forces which seek to avoid unpleasant changes which may accompany moralization (e.g. hedonistic desires). Consistent with this account, Feinberg et al. (2019) found that self-reported hedonic motivations such as the perceived tastiness of meat and seeing meat-eating as a natural part of life, were both negative predictors of moralization.

Future moralization research may also investigate whether cognitive mechanisms may also be applicable to the moralization process. In this regard, one particularly important subset of de-moralization research may be studies investigating various processes by which entities may be viewed as lacking in morally relevant attributes, in order to make actions which cause them harm or issues which negatively affect their interests less morally troublesome. For example, once reminded of past atrocities against an outgroup, members of the perpetrating group have been found to dehumanize members of the outgroup, making their ill-treatment seem less morally troublesome, and thereby restoring psychological equanimity (Castano & Giner-Sorolla, 2006; Vaes, Paladino, & Puvia, 2011). Other work on

phenomena such as the "meat-paradox" (Bastian, Loughnan, Haslam, & Radke, 2012; Loughnan, Haslam & Bastian, 2010), has found that individuals often seek to withdraw their moral concern for certain entities (e.g. Cows or Sheep), when they have made the decision to eat beef or lamb (Buttlar & Walther, 2019). Such research suggests that individuals are motivated to maintain consistency between their moral attitudes and behaviour (Harmon-Jones, Harmon-Jones, & Levy, 2015), however, as far as we are aware, there has not been any corresponding research to date investigating whether a similar process will apply in the context of moralization – that is, how increasing the moral significance of a given action may influence the level of moral significance applied to relevant attitudes and entities.

Conclusion

Over two decades have passed since Rozin (1997) first elucidated the concept of moralization. Since then there have been many studies which have relied on the concept of moralization or have sought to directly study it, yet this brief review has revealed that surprisingly little attention has been paid to the question of how moralization should be operationalized or how various approaches are theoretically linked. In response, we have delineated research examining the moralization of specific actions and compared this to the moralization of attitudes. We have also highlighted the potential for moralization to not only be applied to actions themselves, but also to actors or recipients of those actions. Both people and entities themselves can be considered as more or less morally-relevant, therefore fitting the broad definition of moralization. Our review has also highlighted that although moralization has traditionally been conceptualized as a morally neutral target being newly incorporated within the moral domain, moralization has mostly been studied as incremental increases in the moral importance or wrongness of attitudes or actions that were already considered somewhat morally relevant. In an effort to address this lack of conceptual clarity, we have proposed a new definitional framework for moralization, which makes a distinction

between moral recognition and moral amplification. By taking a broader and more inclusive

approach to the concept of moralization, we have aimed to provide a platform from the

various antecedents, consequences, and processes of moralization can be studied and

understood.

References

Alston, W. P. (1968). Moral Attitudes and Moral Judgments. *Nous, 2*(1), 1-23. doi:10.2307/2214411

Avramova, Y. R., & Inbar, Y. (2013). Emotion and moral judgment. In (Vol. 4, pp. 169-178). Hoboken, USA.

Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. (2012). Don't mind meat? The denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin, 38*(2), 247-256.

Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots: Why we fail to do what's right and what to do about it*: Princeton University Press.

Bloom, P. (2010). How do morals change? *Nature, 464*, 490. doi:10.1038/464490a

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In *Affect, cognition and stereotyping* (pp. 13-37): Elsevier.

Brandt, M. J., Crawford, J. T., & Van Tongeren, D. R. (2017). Worldview Conflict in Daily Life. *Social Psychological and Personality Science*, 1948550617733517.

Brandt, M. J., Wetherell, G., Crawford, J. T., Forgas, J., Jussim, L., & van Lange, P. (2016). Moralization and intolerance of ideological outgroups. *The Social Psychology of Morality*, 239.

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 US presidential election campaign. *Journal of Social and Political Psychology, 3*(2), 211-237.

Buttlar, B., & Walther, E. (2019). Dealing with the meat paradox: Threat leads to moral disengagement from meat consumption. Appetite, 137, 73-80. doi:https://doi.org/10.1016/j.appet.2019.02.017

Castano, E., & Giner-Sorolla, R. (2006). Not quite human: Infrahumanization in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology, 90*(5), 804.

Clifford, S. How Emotional Frames Moralize and Polarize Political Attitudes. *Political Psychology, 0*(0). doi:doi:10.1111/pops.12507

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: a standardized stimulus database of scenarios based on moral foundations theory. *Behavior research methods, 47*(4), 1178-1198. doi:10.3758/s13428-014-0551-2

Cohen, T. R., Montoya, R. M., & Insko, C. A. (2006). Group Morality and Intergroup Relations: Cross-Cultural and Experimental Evidence. *Personality and Social Psychology Bulletin, 32*(11), 1559-1572. doi:10.1177/0146167206291673

Cooper, J. (2007). *Cognitive dissonance : fifty years of a classic theory*: Sage Publications.

Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral Expansiveness: Examining Variability in the Extension of the Moral World. *Journal of Personality and Social Psychology*. doi:10.1037/pspp0000086

Crimston, D., Hornsey, M. J., Bain, P. G., & Bastian, B. (2018). Toward a Psychology of Moral Expansiveness. *Current Directions in Psychological Science, 27*(1), 14-19. doi:10.1177/0963721417730888

Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*(1), 49-65. doi:doi:10.1111/j.2044-835X.1983.tb00543.x

Dewey, J. (1902). The evolutionary method as applied to morality. *The Philosophical Review, 11*(2), 107-124.

Dewey, J., & Tufts, J. H. (1932). Ethics (rev. ed.). *New York: H. Holt and company*.

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology, 44*(4), 1204-1209. doi:10.1016/j.jesp.2008.03.012

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology*.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, Calif Stanford University Press.

Frimer, J. A., Skitka, L. J., & Motyl, M. (2017). Liberals and conservatives are similarly motivated to avoid exposure to one another's opinions. *Journal of Experimental Social Psychology, 72*, 1-12. doi:https://doi.org/10.1016/j.jesp.2017.04.003

Gallup. (2019). Moral Issues. from Gallup Organization https://news.gallup.com/poll/1681/moral-issues.aspx

Gino, F., & Mogilner, C. (2014). Time, money, and morality. *Psychological Science, 25*(2), 414-421.

Goodwin, G. P. (2015). Moral Character in Person Perception. *Current Directions in Psychological Science, 24*(1), 38-44. doi:doi:10.1177/0963721414550709

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition, 106*(3), 1339-1366.

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology, 48*(1), 250-256. doi:https://doi.org/10.1016/j.jesp.2011.08.006

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.

Graham, J., Waytz, A., Meindl, P., Iyer, R., & Young, L. (2017). Centripetal and centrifugal forces in the moral circle: Competing constraints on moral learning. Cognition, 167, 58-65. doi:https://doi.org/10.1016/j.cognition.2016.12.001

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences, 11*(8), 322-323. doi:10.1016/j.tics.2007.06.004

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814-834. doi:10.1037/0033-295X.108.4.814

Haidt, J. (2003). The moral emotions. *Handbook of affective sciences, 11*, 852-870.

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science, 316*(5827), 998-1002. doi:10.1126/science.1137651

Haidt, J., & Graham, J. (2007). When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals may not Recognize. *Social Justice Research, 20*(1), 98-116. doi:10.1007/s11211-007-0034-z

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*(4), 613.

Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An Action-Based Model of Cognitive-Dissonance Processes. *Current Directions in Psychological Science*, 24(3), 184-189. doi:10.1177/0963721414566449.

Horberg, E. J., Oveis, C., & Keltner, D. (2011). Emotions as Moral Amplifiers: An Appraisal Tendency Approach to the Influences of Distinct Emotions upon Moral Judgment. *Emotion Review, 3*(3), 237-244. doi:10.1177/1754073911402384

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology, 97*(6), 963.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: Two faces of moral regulation. *Journal of Personality and Social Psychology, 96*(3), 521.

Johnson, D. J., Wortman, J., Cheung, F., Hein, M., Lucas, R. E., Donnellan, M. B., . . . Narr, R. K. (2016). The Effects of Disgust on Moral Judgments. *Social Psychological and Personality Science, 7*(7), 640-647. doi:10.1177/1948550616654211

Jordan, A. H., & Monin, B. (2008). From sucker to saint: Moralization in response to self-threat. *Psychological Science, 19*(8), 809-815.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-development approach. *Moral development and behavior: Theory research and social issues*, 31-53.

Kohlberg, L. (1984). *The psychology of moral development : the nature and validity of moral stages / Lawrence Kohlberg*. San Francisco: San Francisco : Harper & Row.

Landy, J. F., & Goodwin, G. P. (2015). Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence. *Perspect Psychol Sci, 10*(4), 518-536. doi:10.1177/1745691615583128

Lindström, B., Jangard, S., Selbing, I., & Olsson, A. (2017). The Role of a "Common Is Moral" Heuristic in the Stability and Change of Moral Norms.

Loughnan, S., Haslam, N., & Bastian, B. (2010). The role of meat consumption in the denial of moral status and mind to meat animals. *Appetite*, *55*(1), 156-159.

Lovett, B. J., & Jordan, A. H. (2010). Levels of moralisation: a new conception of moral sensitivity. *Journal of Moral Education, 39*(2), 175-189. doi:10.1080/03057241003754914

Mills, J., & Harmon-Jones, E. (1999). *Cognitive dissonance: progress on a pivotal theory in social psychology* (Vol. 1st). Washington, DC American Psychological Association.

Monroe, A. E., Dillon, K. D., Guglielmo, S., & Baumeister, R. F. (2018). It's not what you do, but what everyone else does: On the role of descriptive norms and subjectivism in moral judgment. *Journal of Experimental Social Psychology, 77*, 1-10. doi:https://doi.org/10.1016/j.jesp.2018.03.010

Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature Human Behaviour*, 1.

Morris, M. W., & Liu, Z. (2015). Psychological Functions of Subjective Norms:Reference Groups, Moralization, Adherence, and Defiance. *Journal of Cross-Cultural Psychology, 46*(10), 1279-1287. doi:10.1177/0022022115612161

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child development*, 400-407.

Pizarro, D., Inbar, Y., & Helion, C. (2011). On Disgust and Moral Judgment. *Emotion Review, 3*(3), 267-268. doi:10.1177/1754073911402394

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*(1), 57-75. doi:10.1037/a0021867

Roberts, S. O., Guo, C., Ho, A. K., & Gelman, S. A. (2018). Children's descriptive-to-prescriptive tendency replicates (and varies) cross-culturally: Evidence from China. *Journal of Experimental Child Psychology, 165*, 148-160. doi:https://doi.org/10.1016/j.jecp.2017.03.018

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2017). Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive. *Journal of Experimental Child Psychology, 158*, 19-31. doi:https://doi.org/10.1016/j.jecp.2016.11.013

Rottman, J., & Kelemen, D. (2012). Aliens behaving badly: Children's acquisition of novel purity-based morals. *Cognition, 124*(3), 356-360. doi:https://doi.org/10.1016/j.cognition.2012.06.001

Rottman, J., Young, L., & Kelemen, D. (2017). The impact of testimony on children's moralization of novel actions. *Emotion, 17*(5), 811.

Royzman, E. B., Landy, J. F., & Goodwin, G. P. (2014). Are good reasoners more incest-friendly? Trait cognitive reflection predicts selective moralization in a sample of American adults. *Judgment and Decision Making, 9*(3), 175.

Rozin, P. (1997). Moralization. In A. M. Brandt & P. Rozin (Eds.), *Morality and Health*. Florence, UNITED KINGDOM: Routledge.

Rozin, P. (1999). The process of moralization. *Psychological Science, 10*(3), 218-221.

Rozin, P., & Singh, L. (1999). The moralization of cigarette smoking in the United States. *Journal of Consumer Psychology, 8*(3), 321-337.

Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a vegetarian: The transformation of preferences into values and the recruitment of disgust. *Psychological Science, 8*(2), 67-73.

Ryan, T. J. (2014). Reconsidering Moral Issues in Politics. *The Journal of Politics, 76*(2), 380-397. doi:10.1017/s0022381613001357

Ryan, T. J. (2017). No compromise: Political consequences of moralized attitudes. *American Journal of Political Science, 61*(2), 409-423.

Schaller, M., & Park, J. H. (2011). The behavioral immune system (and why it matters). *Current Directions in Psychological Science, 20*(2), 99-103.

Schein, C., & Gray, K. (2014). The Prototype Model of Blame: Freeing Moral Cognition From Linearity and Little Boxes. *Psychological Inquiry, 25*(2), 236-240. doi:10.1080/1047840X.2014.901903

Schein, C., & Gray, K. (2016). Moralization and Harmification: The Dyadic Loop Explains How the Innocuous Becomes Harmful and Wrong. *Psychological Inquiry, 27*(1), 62-65. doi:10.1080/1047840X.2016.1111121

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review, 22*(1), 32-70.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as Embodied Moral Judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096-1109. doi:10.1177/0146167208317771

Sezer, O., Gino, F., & Bazerman, M. H. (2015). Ethical blind spots: Explaining unintentional unethical behavior. *Current Opinion in Psychology, 6*, 77-81.

Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology, 102*(6), 1164.

Skitka, L. J. (2010). The Psychology of Moral Conviction. *Social and Personality Psychology Compass, 4*(4), 267-281. doi:10.1111/j.1751-9004.2010.00254.x

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral Conviction: Another Contributor to Attitude Strength or Something More? *Journal of Personality and Social Psychology, 88*(6), 895-917. doi:10.1037/0022-3514.88.6.895

Skitka, L. J., Wisneski, D. C., & Brandt, M. J. (2017). Attitude moralization: Probably not intuitive or rooted in perceptions of harm. *Current Directions in Psychological Science*, 0963721417727861.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annu. Rev. Psychol., 58*, 345-372.

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research, 17*(2), 223-236.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behavior relationship: A role for group identification. *Personality and Social Psychology Bulletin, 22*(8), 776-793.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440-463. doi:10.1037/a0018963

Turiel, E. (1983). *The development of social knowledge: Morality and convention*: Cambridge University Press.

Turner-Zwinkels, F. M., Zomeren, M., & Postmes, T. (2017). The moral dimension of politicized identity: Exploring identity content during the 2012 Presidential Elections in the USA. *British Journal of Social Psychology, 56*(2), 416-436.

Vaes, J., Paladino, P., & Puvia, E. (2011). Are sexualized women complete human beings? Why men and women dehumanize sexually objectified women. *European Journal of Social Psychology, 41*(6), 774-785.

van Bavel, J. J., Packer, D. J., Haas, I. J., & Cunningham, W. A. (2012). The Importance of Moral Construal: Moral versus Non-Moral Construal Elicits Faster, More Extreme, Universal Evaluations of the Same Actions. *PLoS ONE, 7*(11), e48693. doi:10.1371/journal.pone.0048693

van Bavel, J. J., Packer, D. J., Haas, I. J., & Cunningham, W. A. (2012). The Importance of Moral Construal: Moral versus Non-Moral Construal Elicits Faster, More Extreme, Universal Evaluations of the Same Actions. *PLoS ONE, 7*(11), e48693. doi:10.1371/journal.pone.0048693

van Zomeren, M. (2013). Four Core Social-Psychological Motivations to Undertake Collective Action. *Social and Personality Psychology Compass, 7*(6), 378-388. doi:10.1111/spc3.12031

van Zomeren, M. (2016). Building a Tower of Babel? Integrating Core Motivations and Features of Social Structure into the Political Psychology of Political Action. *Political Psychology, 37*, 87-114. doi:10.1111/pops.12322

van Zomeren, M., Postmes, T., Spears, R., & Bettache, K. (2011). Can moral convictions motivate the advantaged to challenge social inequality?: Extending the social identity model of collective action. *Group Processes & Intergroup Relations, 14*(5), 735-753. doi:10.1177/1368430210395637

Vauclair, C. M., & Fischer, R. (2011). Do cultural values predict individuals' moral attitudes? A cross-cultural multilevel approach. *European Journal of Social Psychology*, *41*(5), 645-657.

Wheatley, T., & Haidt, J. (2005). Research Report: Hypnotic Disgust Makes Moral Judgments More Severe. *Psychological Science, 16*(10), 780-784. doi:10.1111/j.1467-9280.2005.01614.x

Wisneski, D. C., & Skitka, L. J. (2016). Moralization Through Moral Shock. *Personality and Social Psychology Bulletin, 43*(2), 139-150. doi:10.1177/0146167216676479

2

Chapter 2
Toward a Functional Threat-Based Account of
Moralization

Chapter 2 Toward a Functional Action-Based Moralization Account

2.1. Preface

This chapter presents an article proposing a novel theoretical framework of moralization that focuses on how the process is functional for the individual. The manuscript in this chapter is currently in being prepared to submit for publication.

In Chapter 1, we developed a taxonomic framework aimed to capture the functional dimensions of variability present in the current literature on moralization. In doing so, we proposed that moralization encompasses increases in moral significance attributed to actions, attitudes and entities. We also refined the definition of moralization to clarify that it includes both the categorical movement from moral to non-moral ("moral recognition"), and a continuous movement in the extremity of moral significance ("moral recognition").

If the goal of Chapter 1 was to clarify "what" moralization is, in Chapter 2, we use this definition of moralization as a starting point for developing a functional account of moralization, which forms the theoretical framework for the empirical work presented in the remainder of the thesis. Broadly, the objective of the framework – which we call the action-based moralization account – is to better understand the functional motivators of moralization from the perspective of the individual. In other words, the remainder of the thesis is focused around the question of "when" and "why" moralization might occur.

The objective of the action-based moralization account is to identify the functional dimensions of variability occurring within the individual that can explain when they are likely to attribute greater moral significance to actions or attitudes, and why it is adaptive for them to do so. Such an account addresses a current gap in theorising in moralization research that has predominantly focused on the either the proximal mechanisms of moralization (i.e. *how* it occurs), or on the functional benefits of moralization for groups.

To foreshadow the major premises of the action-based moralization account, we propose that moralization is a functional response to threat-appraisal that actives schemas of cognitions and behaviours which allow the individual to resolve uncertainties regarding current and future action. Specifically, we argue that, in the short-term, moralization is a functional response to threat as it allows the individual to make quick and efficient decisions by splitting potential choice alternatives and allows them to maintain certainty about their beliefs by creating a source of epistemic authority that is resistant to updating based on new information. In the long-term, we argue that moralization facilitates the individual's ability to coordinate with other belief sharing individuals to respond to new threats.

The manuscript in this chapter is followed by a summary of the program of empirical research was developed in this thesis as a means of evaluating the initial feasibility of the action-based moralization account.

2.2.   Manuscript

Towards an Action-based Moralization Account

Since the very beginning of evolutionary theory, it has been argued that morality is an essential ingredient for human social cohesion and coordination (Darwin, 1871). Indeed traditional evolutionary and lay understandings of morality treat it as a universal motivator of altruism (FeldmanHall et al., 2018), fairness (de Waal & Waal, 1996), and a signal of trustworthiness (Goodwin, 2015). However, the past 20 years of moral psychology research has revealed that what is subjectively considered to be morally relevant can vary substantially between individuals of a given society (Ryan, 2014). Perhaps more surprising has been the insight that certain types of stimuli (e.g. images that induce strong negative emotions) can impact the degree to which a given person might consider certain attitudes to be a matter of moral conviction, or certain actions to be morally wrong (Haidt, 2001; Skitka et al., 2021). For the purposes of this paper, I refer to this phenomenon of within-person variation in the degree to which moral relevance is attributed to beliefs or judgments as *moralization* (Rhee et al., 2019; Rozin, 1999).

The insight that people can vary in the degree to which they moralize is meaningful given growing evidence that people experience moralized beliefs in a way that is psychologically distinct from other types of beliefs (Skitka et al., 2021). Moralized beliefs are perceived as akin to objective facts (Goodwin & Darley, 2008, 2012) and universally prescriptive regardless of context (Skitka et al., 2005; van Bavel et al., 2012). People also take active measures to maintain the stability of their moralized beliefs, such as distancing themselves from people who hold contrary views (Cole Wright et al., 2008; Luttrell & Togans, 2021; Skitka et al., 2005), and even adjusting their perspectives on fact-based arguments to be more consistent with their moralized positions (Gampa et al., 2019; Liu & Ditto, 2013). Concerningly, as research into the distinct psychological qualities of subjective

moralized beliefs has developed, so too has evidence of the potentially harmful social and political consequences arising from the unwavering nature of moralized beliefs (Kovacheff et al., 2018). Indeed, research suggests that moralization of beliefs can increase individuals' tendency to spread misinformation (Ecker et al., 2022; Effron & Raj, 2020), become more intolerant of outgroups (Ballone et al., 2023; Brandt et al., 2016), and support the use of violence against groups with opposing moral beliefs (Mooijman et al., 2018).

The revelations on the potentially harmful outcomes of moralization clearly highlight the need for a comprehensive understanding of the process that will allow us to better predict the real-world conditions that increase people's tendency to moralize, and when such moralization is likely to produce undesirable social or political outcomes. Such an understanding requires interrogation of not only of the mechanisms of underlying moralization (the "*how*"), but also of the functional or adaptive reasons that individuals moralize in a given context or environment (the "*why*"). In this paper I argue that current theorising on subjective moralization at the individual-level has focused predominantly on the *how*-question but has not engaged deeply with the *why*-question. I then advocate for a functional approach to understanding moralization and propose a new functional and integrative theoretical framework called the action-based moralization account. Through this new account, I propose an understanding of moralization that revolves around facilitating the motivation to engage in quick and decisive action. Specifically, I argue that moralization activates schemas of cognitions and behaviours that function to quickly overcome uncertainty under situations of high threat-appraisal thus enabling the moralizer to act more decisively in future contexts, while simultaneously preparing the moralizer to engage in effective group-based coordination against the threat.

A brief overview of current theories on subjective moralization

The inception of moralization research was firmly rooted in the foundational belief that morality, as a construct, described an objective set of attitudes and behaviours (for review see: Rhee et al., 2019). Naturally, this meant early study of moralization primarily centred on how individuals acquired varying degrees of ability to engage in complex moral reasoning (e.g. Kohlberg, 1976). However, the trajectory of moralization research began to change alongside an increasing awareness that substantial variance could exist among individuals in their perceptions of the types of attitudes or actions that were morally wrong. The seminal work of Rozin and colleagues, who investigated changes in moral attitudes toward smoking and vegetarianism (Rozin, 1999; Rozin et al., 1999), marked a shift in the focus of moral theorising from trying to understand how people achieve *objective* moral standards, towards understanding the factors that led to changes in *subjective* moral beliefs (for review see: Rhee et al., 2019).

Proximal Accounts and Subjective Approaches to Moralization

The initial theoretical focus of advocates of subjective approaches to moralization was to develop a framework for understanding the parameters of what individuals consider to be morally relevant (Haidt, 2007). In order to make the case that people do not determine the moral relevance of their beliefs by referring to a universal set of moral standards, it was necessary for theories of subjective morality to define how people recognise the moral relevance of their attitudes or judgments (Greene & Haidt, 2002; Haidt, 2001, 2007). Naturally, this meant that early subjective theories of moralization focused on identifying psychological mechanisms that were reliably *proximal* to moral judgments or attitudes and could therefore be seen as a psychological cue for moral relevance (see Figure 1).

An early example of a *proximal account* of moralization is the Social Intuitionist Theory of morality (SIT: Haidt, 2001). According to SIT, people's moral beliefs are largely acquired and continuously updated implicitly through social and cultural interactions (Haidt,

2007). Crucially, SIT argues that most moral beliefs manifest through implicit cues that trigger the experience of morally relevant emotions (particularly disgust), which in turn allow us to intuit moral relevance (Schnall et al., 2008). Another influential proximal account of morality is the theory of dyadic morality (TDM: (Schein & Gray, 2016), which centres on the premise that morality is defined by the concept of harm. Specifically, TDM posits that moral beliefs are centred around a harm-based cognitive template, which is inherently dyadic, consisting of an agent and a patient. The theory suggests that perceptions of harm lead to moral condemnation (dyadic comparison), and judgments of immorality lead to perceptions of harm (dyadic completion). This creates a feedback cycle, termed the "dyadic loop," which contributes to the moral acquisition and polarization over time.

Figure 1

Theorising on moralization can happen at multiple levels. Most accounts of moralization to date have focused on answering research questions that arise when trying to explain the proximal mechanisms of moralization.

	Individual-Level		Group-Level
	Mechanism	Function	
	Proximal Mechanisms	Distal Motivations	Group-Level Functions
Focal research questions	• How does moralization occur? • What are the cognitive or emotional mechanisms of moralization?	• How does a person's environment impact their tendency to moralize? • Why are the outcomes of moralization functional for the individual?	• How does moralization facilitate group cooperation?
Applicable theories	• Social intuitionist theory (Haidt, 2001; 2007) • Theory of dyadic morality (Schein & Gray, 2016)	• Action-based moralization account	• Dynamic coordination account (Descioli & Kurzban, 2012) • Morality as cooperation account (Curry, 2016)

Limitations of proximal accounts

Theoretical frameworks that identify the *proximal* mechanisms of moralization have played a critical role in enabling meaningful study of subjective morality, and are necessary foundations for the development of psychological interventions targeted at shifting specifical moral beliefs (Feinberg et al., 2019). Indeed, the empirical evidence suggests that both harm-related cognitions and moral emotions may be relevant "push" appraisal factors (Feinberg et al., 2019) that make it more likely that an individual will moralize a relevant belief after being presented with a morally evocative stimulus (Feinberg et al., 2019; Rhee et al., 2019; Skitka et al., 2021).

However, the limitation of proximal analyses in psychology is that they often risk over-simplification, potentially reducing complex behaviours or cognitions to one or two psychological processes. Further, because proximal accounts primarily address immediate or recent antecedents of psychological phenomena, they do not readily generate testable explanations for how the behaviour or cognition in question is likely to be impacted by factors in a person's natural environment (Fiedler, 2014; Nelissen et al., 2013). Reflecting these limitations, both harm and disgust have proved to be inconsistent predictors of moralization in longitudinal or field studies. For example, (Brandt et al., 2015) found that disgust but not perceived harm was a significant predictor of moralization around the 2012 US presidential election. Conversely, (Wisneski et al., 2020) found that only harm – but not disgust – was a predictor of moralization of participants' attitudes around an ongoing water contamination crisis taking place in the US city of Flint, Michigan.

More importantly for the present purposes, current proximal theories of moralization such as SIT and TDM do not readily account for how the process of moralization is related to empirical outcomes of moralization at the individual level, or indeed why it is functional for the individual to engage in moralization. This means that these theories have limited utility

when it comes to predicting the psychological outcomes of moralization that have been

brought to light in more recent research (Skitka et al., 2021).

Toward a functional account of moralization

Moral concepts can't be formed by abstraction from any empirical

knowledge or, therefore, from anything contingent; that this purity or non-

empiricalness of origin is what gives them the dignity of serving as supreme

practical principles; that any addition of something empirical takes away

just that much of their influence and of the unqualified worth of actions

performed in accordance with them – Kant, 1785 p. 17

At the most general level, a functional account is an explanation of psychological

phenomena through the perspective of how they are adaptions to the problem of social and

physical survival (Keltner & Gross, 1999). Compared to proximal approaches which mainly

focus on identifying immediate precedents of the psychological phenomenon to be explained,

functional accounts approach theorising in terms of identifying the features of a person's

physical or social context that the phenomenon helps the person overcome (Fiedler, 2014).

This integration of physical or social factors in understanding *when* and *why* a person might

engage in certain behaviours or cognitions means that functional accounts can more readily

generate predictions about *when* and *why* those behaviours or cognitions are likely to arise in

other contexts (Fiedler, 2014; Scott-Phillips et al., 2011).

In the case of moralization, a functional approach would seek to understand two key

elements: (1) how the outcomes of moralization fill or address a deficit or need for the

individual within their physical or social environment; and (2) when and how the process of

moralization is activated to respond to that need. In this section, I attempt to answer these two

questions by proposing a novel functional framework called the action-based moralization

account. The action-based moralization account revolves around facilitating the motivation to

engage in quick and decisive action. Specifically, I argue that moralization activates schemas of cognitions and behaviours that function to quickly overcome uncertainty under situations of high threat-appraisal, and enables the moralizer to act more decisively in future contexts.

Empirical outcomes of moralization

To engage in a thorough analysis of how the outcomes of moralization may fill or address a deficit or need for the individual within their physical or social environment, it is important to start with a discussion of the empirical evidence to date. The movement toward treating variance in subjective moral belief as a meaningful unit of moral psychology research has also led to new insights into how moral beliefs might be different from non-moral beliefs. The biggest contribution to this direction of inquiry has come from the moral conviction program (Skitka et al., 2005), which conceptualizes variance in subjective moralization in terms of how strongly participants see an issue as connected to their core moral beliefs and convictions (Skitka et al., 2021). The moral conviction approach allows for observation of variance in subjective moralization without referring to participants' specific position on any given issue (e.g. whether they are "pro-life" or "pro-choice" on the topic of abortion), which in turn allows researchers to more reliably identify patterns of cognition and behaviour that are specifically relevant to variance in moralization regardless of the specific content of the moralizing individual's beliefs.

Belief Insulation

A well-established empirical characteristic of moralization is that that people become highly motivated to maintain the stability of their position on moralized beliefs (Cole Wright et al., 2008; Luttrell & Togans, 2021; Skitka et al., 2005). This motivation manifests in both psychological and physical distancing from people who hold beliefs that conflict with the individual's moralized position. People with stronger moral convictions are more likely to indicate that they would not be happy to have someone who disagreed with their beliefs in

their social circle (e.g. as a room mate: Skitka et al., 2005). When given the choice to configure the layout of chairs in a room, people with strong moral convictions are also more likely to sit further away from others who they believe to have opposing beliefs (Skitka et al., 2005).

Belief Factualization

An intriguing characteristic category of empirical of moralization outcomes is the way the process appears to influence perceptions of factual information. There are two ways in which moralized beliefs appear to be *factualized* (Ditto & Liu, 2016). The first is that moralized beliefs are themselves perceived to be akin to objective fact. More specifically, individuals are likely to perceive morally laden statements as having clear true or false answers as opposed to being a matter of opinion or attitudes (Goodwin & Darley, 2008), and are more likely to interpret disagreement with their position as clearly mistaken (Goodwin & Darley, 2008, 2012). Second, individuals also factualize their moralized beliefs by using them as a basis for generating new factual beliefs and processing incoming information. For example, people who have strongly moralized their position against the use of capital punishment are more likely to believe the factual claim that the cost of keeping an inmate on death row is unjustifiably high compared to the benefit (Liu & Ditto, 2013).

Belief-based Mobilization

Moral beliefs are experienced as inherently obligatory and motivating (Brandt et al., 2015; Skitka et al., 2021). They compel individuals to act in accordance with their moral beliefs, often resulting in increased political engagement and activism (Louis et al., 2022; van Zomeren et al., 2012; Zaal et al., 2011).

Moral Convergence Amplifies the Characteristics of Moralization

Moral convergence refers to the perception by a moralizing individual that a majority of other people in a referent group have also moralized the same belief (Mooijman et al.,

2018). While the empirical research looking at the impact of moral convergence is small compared to the other outcomes of moralization described in this section, there is some empirical evidence to suggest that perceptions of moral convergence amplify, or sometimes even completely moderates, the psychological characteristics of moralization (Mooijman et al., 2018). For example, Goodwin & Darley (2012) found that participants who were told their moral beliefs aligned with the majority view were more likely to perceive the belief as akin to objective fact compared to those who were told that they did not align with the majority view. More recently, across three studies, Mooijman et al. (2017) consistently found that the relationship between belief moralization and support for violence against belief opponents was fully moderated by perceived moral convergence.

Making Sense of Empirical Outcomes: The Action-based Moralization Account

Collectively, the empirical evidence suggests that moralization produces distinct psychological changes in the way they interact with their moralized beliefs. Specifically, when beliefs are moralized, individuals appear more motivated to *insulate* those beliefs from outside influence, are more likely to *factualize* their beliefs, and are more likely to be mobilized to take action in support of their beliefs. The core premise of the action-based account is that these empirical outcomes of moralization represents activation of a range of cognitive schemas and behaviours that facilitate the resolution of anxious-uncertainty under conditions of perceived threat. In particular, the action-based account identifies two mechanisms by which moralization may function to resolve uncertainty by: (1) facilitating efficient action by splitting of belief alternatives; and (2) creating the basis for an internalised epistemic standard for processing new information. Finally, the action-based account links these individual level processes to group dynamics by proposing that the formation of highly stable beliefs at the individual level facilitates the formation and maintenance of belief-based groups.

Figure 2

A conceptual diagram of the action-based moralization account
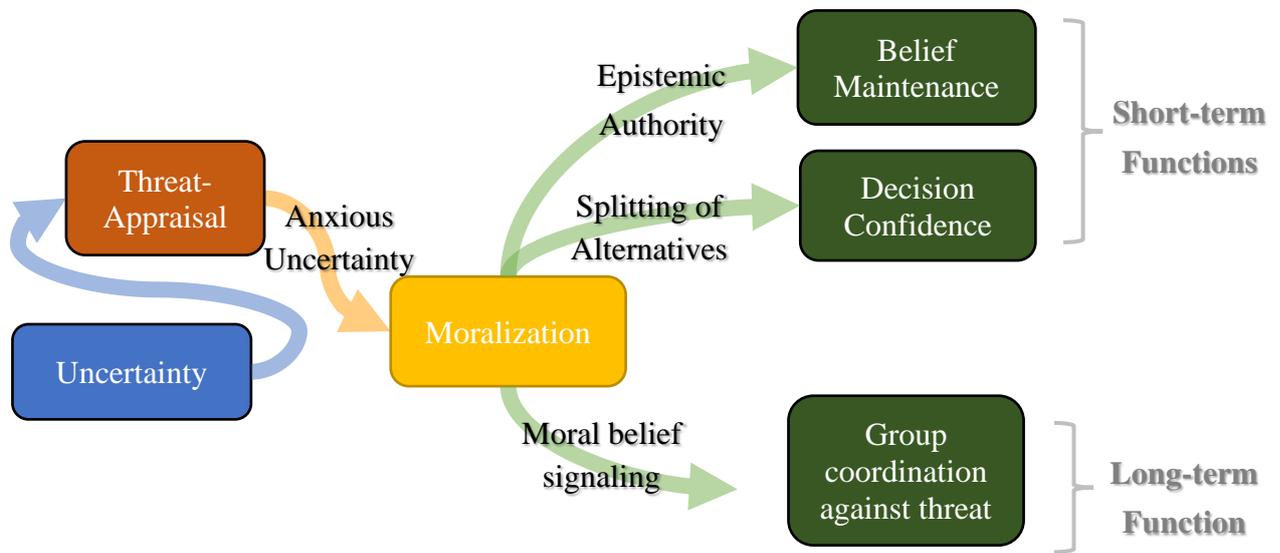
Decision-Making under Uncertainty

Uncertainty is an unavoidable part of our daily interactions with the social and physical environment around us. The complexity of the world that I are required to navigate means that our actions are often based on predictions about what the likely outcome of our choices will be, even where I have little experience or information available to support such predictions (Tversky & Kahneman, 1974). The cognitive tools that humans apply to take action in the face of such limited information has been a topic of interest for cognitive, behavioural and social psychology researchers (Haas, 2016; Kruglanski, 1989; Oeberst & Imhoff, 2023). Such research has found that humans will often apply a range of cognitive shortcuts and heuristics to make decisions in cases of uncertainty where they do not have access to all relevant information, or have limited cognitive processing capacity available to thoroughly evaluate the available information (Oeberst & Imhoff, 2023; Tversky & Kahneman, 1974).

The particular strategies I apply to make decisions in the face of uncertainty will often depend on the context in which the decisions are made, and the particular motivation that is

salient at the time (Kruglanski et al., 2010). For example, when the motivation to gain an accurate world view is salient, I may respond to uncertainty by engaging in further information exploration in order to obtain better domain-relevant information (Gershman, 2019). However, there may be other situations where I are motivated to quickly resolve uncertainty so that I can immediately form a decision, impression or belief (Dechesne & Kruglanski, 2004). In such situations, the optimum strategy may in fact be to interpret new incoming information in a way that confirms our existing beliefs (Peters, 2022), or to seek what has been termed non-specific cognitive closure (Dechesne & Kruglanski, 2004) – that is, to accept any information system as a guide to decision-making as long as it is capable of providing cognitive closure (Kruglanski & Webster, 1996).

Threat-Appraisal and Uncertainty Motivations

A particularly powerful contextual elicitor of the motivation for non-specific closure is the appraisal of threats in one's environment (Dechesne & Kruglanski, 2004; Haas, 2016; Haas & Cunningham, 2014; Jonas et al., 2014; McGregor et al., 2010). According to Gray & McNaughton's (2000) highly influential model of approach and avoidance, where an individual detects threat-relevant cues, a series of neuropsychological processes are activated *via* the behavioural inhibition system (Gray, 1975) that increases their vigilance to other potential dangers, and alerts them of the need to disengage their current goals (McGregor et al., 2010; McNaughton & Corr, 2004; Tomaka et al., 1997). This manifests in a highly aversive state of anxious-uncertainty which impedes any future goal-driven behaviour until the uncertainty is resolved (McGregor et al., 2010).

According to lay epistemics theory (Kruglanski, 1990) such aversive uncertainty states are particularly likely to generate motivations for non-specific closure (Kruglanski & Webster, 1996). In other words, individuals experiencing threat-induced anxious-uncertainty will be particularly motivated to adopt broad knowledge systems that allow them to quickly

resolve their current state of anxiety, and also allow them to quickly resolve future uncertainty, thus facilitating engagement in decisive future action (Dechesne & Kruglanski, 2004; Kruglanski & Webster, 1996).

How moralization resolves action-based uncertainty: *Moralization as spreading of belief alternatives*

The action-based moralization account proposes that moralization can be conceptualized as a process of splitting belief alternatives. In other words, moralization represents the selective privileging of an action-relevant belief, and active diminishing of other beliefs that might conflict with the moralized belief.

One of the most well-studied cognitive tools that people employ to reduce their uncertainty about future decisions is cognitive dissonance reduction (Festinger, 1957; Harmon-Jones & Mills, 2019). The notion that barriers to future action generate a state of cognitive discomfort is supported by recent advancements in the study of cognitive dissonance. Cognitive dissonance describes a situation of discomfort caused when an individual has two or more related cognitions which are inconsistent with each other (Festinger, 1957). According to Festinger (1957), this state of discomfort caused by a "generative cognition" prompts individuals to engage in cognitive work aimed at reducing dissonant cognitions and increasing consistent ones. Recent developments in neurobiological perspectives have led to the development of an "action-based" model of dissonance (E. Harmon-Jones, Harmon-Jones, & Levy, 2015). The fundamental premise of the action-based model is that a critical function of cognition is to drive behaviour. This means that where related cognitions conflict, further action becomes difficult until the conflict is minimised or resolved. Thus, the action-based model conceptualises dissonance reduction as a means of assisting in the implementation of actions once a decision has been made to engage in them.

The key insight provided by the action-based model is that it identifies cognitive dissonance as a means of facilitating action. In other words, where a person is motivated to take a course of action, cognitions that are congruent with enabling that action will prevail following dissonance resolution (Harmon-Jones et al., 2015). The action-based moralization account proposes that moralization is a process that is very much analogous to the spreading of alternatives (see *Figure 3*). To elaborate, the first premise of the action-based moralization account is that both *belief factualization* and *belief insulation* outcomes of moralization can be understood as the result of a psychological process that reduces uncertainty about future action by creating separation between the moralized belief and other belief alternatives. Put another way, moralization provides the individual with greater certainty about their future actions by elevating the beliefs that support their decision, and actively avoiding beliefs that may support an alternative course of action.

Figure 3

Conflicting beliefs can generate uncertainty around the action that an individual should take (Panel A). Moralization reduces this action-based uncertainty through a splitting of alternative beliefs, whereby the moralized belief is elevated and alternative beliefs are actively derogated (Panel B)
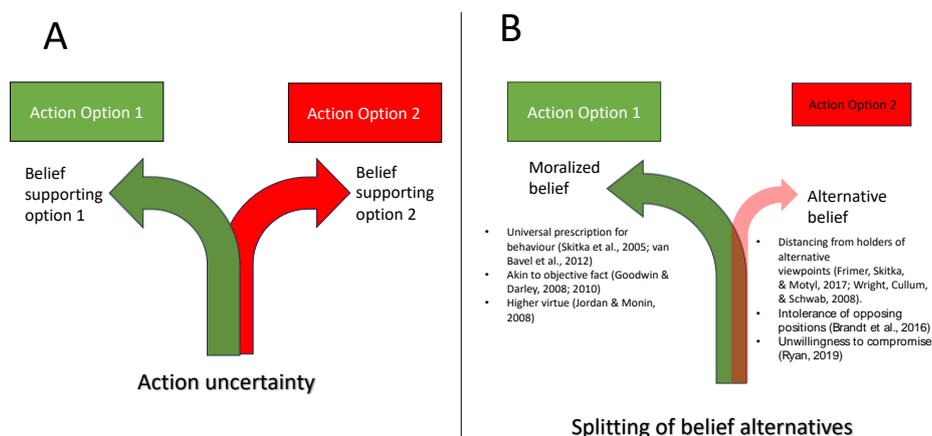
To illustrate this process, let's consider the example of belief in anthropogenic climate change. Long term climate modelling is a highly technical and notoriously difficult field of

science that requires representation of earth system processes in mathematical equations. From the individual's perspective, the informational uncertainty regarding anthropogenic climate change presents an uncomfortable action-based dilemma. On one hand, if climate change is indeed caused by human activity, the failure to take decisive action to reduce carbon emissions may result in catastrophic outcomes for humanity. On the other, if climate change is not caused by human activity, the individual may incur significant cost on climate mitigation efforts that ultimately have little importance. Regardless of the course of action the individual decides to take, moralization of their chosen position provides them with a basis for action that does not require them to continuously weight the relative merit of alternatives. Instead, alternative positions are ruled out *a priori*, by the fact that they are inconsistent with the moralized position.

Consistent with this account of moralization, converging research indicates that that moral salience does indeed increase the speed and efficiency of decision-making. For example, Van Bavel et al. (2012) assigned participants to rate the same set of actions either in moral terms (i.e. morally good vs. morally bad) or in pragmatic terms (i.e. personally good vs. personally bad). Consistent with the action-based splitting of alternatives perspective, it was found that participants who engaged in moral evaluations tended to be more extreme in their ratings, and also rated the same actions as more universal (i.e. universally prohibited vs. universally good) than participants who engaged in pragmatic evaluations. Further, supporting the notion that moralization facilitates efficient decision-making, it was found that participants in the moral evaluation condition made faster decisions regarding the universality of the actions compared to those in the pragmatic condition. Similarly, Koukachi et al. (2018) found that participants devoted less time and less visual attention to the alternative of two choices where they perceived one of the options as being clearly morally relevant.

Moralization as Source of Epistemic Authority

An important aspect of moralization is how this process allows the individual moralizer to continuously maintain low levels of uncertainty in the face of new information. In this regard, the action-based moralization account proposes that moralization allows individuals to maintain low levels of uncertainty by ascribing *epistemic authority* to the moralized belief. In other words, through the splitting of belief alternatives, moralization also activates a strong tendency toward confirmation bias (Mercier & Sperber, 2011; Peters, 2022) where the moralized prior belief influences the way individuals seek and interpret new information.

According to lay epistemics theory, one source of evidence that I can use in our evaluation of new information is an epistemic authority – that is, a source that I rely on to acquire knowledge (Kruglanski, 1989). Importantly, within the lay epistemics theory framework, the epistemic authority of a source refers to the degree to which an individual is willing to accept it as a source of information, rather than whether it possesses objective characteristics (e.g. expertise) that imply credibility (Kruglanski et al., 2005). This means that the level of epistemic authority ascribed to a given source will depend largely on the attributes, motivations and circumstances of the individual seeking information. Importantly, lay epistemics theory suggests that the characteristics of the epistemic authority that an individual chooses to adopt will be driven by their epistemic motivation. For example, someone with a strong motivation for non-specific cognitive closure may attribute a high degree of general epistemic authority to a single source, meaning that their evaluation of information in various domains are influenced by the same epistemic authority (Kruglanski et al., 2005; Kruglanski & Webster, 1996).

Applying these insights to the action-based account of moralization, I propose that moralization also facilitates the maintenance of low uncertainty and high stability in beliefs by ascribing epistemic authority to the moralized beliefs, which in turn allows the individuals

to judge the factual-value of new information based on their consistency with their internalized moral beliefs. Put another way, the action-based moralization account explains the empirical outcome of factualization following moralization as individuals ascribing epistemic value to their moralized belief, which in turn supports a strong tendency toward confirmation bias such that they only seek out belief-consistent information, or interpret new information as being consistent with their moralized beliefs (Peters, 2022; Yoder & Decety, 2022).

This formulation of moralization also can be illustratively expressed using a Bayesian framework. According to Bayes' theorem, optimal belief updating involves three variables: the probability of a prior belief after observing new data is expressed as $p(\text{H}|\text{D})$ as proportional to product of the likelihood ratio of observing the data if the person's prior belief is true $\frac{p(\text{D}|\text{H})}{p(\text{D})}$, and the probability of the prior belief $p(\text{H})$. According to the action-based moralization account, the effect of moralization is to disruption this process of Bayesian belief updating by: (1) increase the individual's subjective perception regarding probability of their prior beliefs, and (2) making the prior belief an epistemic yardstick for evaluating the likelihood of the new information (for illustration see *Figure 3*). In other words, moralization insulates beliefs from any uncertainty arising from new information by making interpretation of the new information itself dependent on consistency with the moralized belief (Oeberst & Imhoff, 2023).

Figure 4

Illustration of how moralization can interrupt optimal Bayesian belief updating through selective seeking and interpretation of belief-congruent information.
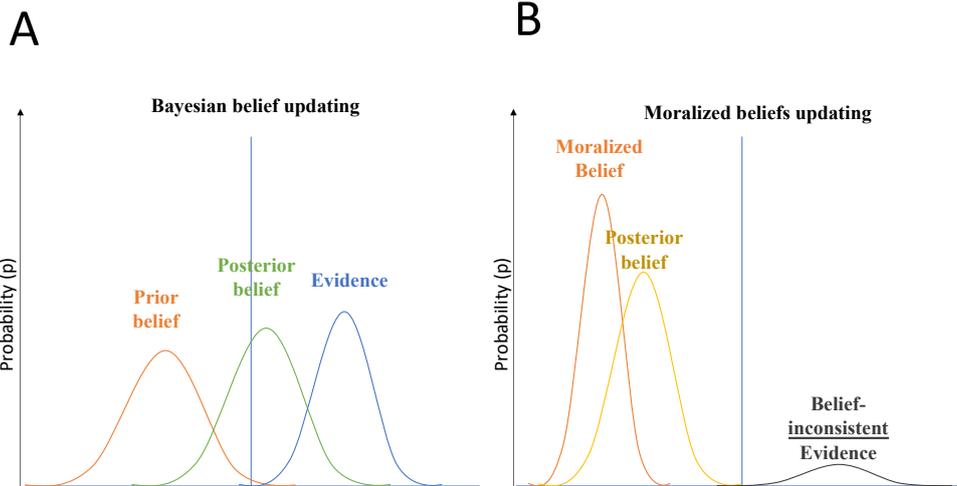
Again, let's return to the example of belief in anthropogenic climate change. Someone who has moralized their position on anthropogenic climate change may quickly experience action-based uncertainty again if they encounter new belief-disconfirming information. However, this ongoing uncertainty can be avoided entirely if the moralized belief (either believing or not believing in anthropogenic climate change) can itself become the indicator of whether the new information is worth engaging with.

Action-based Moralization and Group-based Coordination: An Individual-level Process Leading to a Group-based solution

Thus far, I have proposed that subjective moralization is an adaptive process for the individual insofar as it radically reduces their action-based uncertainty in the face of threat. However, this picture alone is an incomplete account of why moralization is a functional response to threat-appraisal. Indeed, while action-based moralization may function to reduce the individual's internal state of aversive uncertainty, it does not necessarily address the source of the threat appraisal itself. Thus, the third and final proposal of the action-based uncertainty account is that the ultimate function of moralization is to facilitate the individual's ability to engage in group-based cooperation against the threat. Critically, the action-based moralization proposes that individuals will be more likely to moralize to reduce

uncertainty in the presence of threat, precisely because it increases their ability to coordinate with other group members. In other words, while the short term function of moralization may be to reduce immediate uncertainty about their own actions, the long term function of moralization in response to threat is to facilitate coordination with allies who hold the same beliefs (Petersen, 2013).

Threat-appraisal can be conceptualised as an individual's subjective perception that they are faced with a situation that demands action that they do not have the personal resources to deal with (Tomaka et al., 1997). In such circumstances, a powerful strategy that individuals often adopt deal with the situation is to increase their access to social resources by securing the cooperation of others (Blascovich, 2008; Meijen et al., 2020). Recent evolutionary theories of morality suggest that an important group-based function of morality is to act as a dynamic social cue for individual agents to strategically coordinate in the event of conflict (DeScioli & Kurzban, 2013). According to this dynamic coordination account of morality, the key social function of morality is to enable individual agents within a society to publicly signal the opportunity for coordination against a shared opponent.

Integrating the dynamic coordination account of morality with the action-based moralization model, I propose that the individual and group-level functions of moralization intersect through the formation of update resistant and stable beliefs. Specifically, I propose that the stability of the contents of moralized beliefs can facilitate efficient communication between multiple agents by making it easier for each individual agent to more efficiently identify others with shared beliefs (Banisch & Olbrich, 2018). Unlike coordination based on kinship or pre-existing social alliances, a key challenge of coordinating collective action on the basis of morality is identifying others with the same moralized beliefs. To this end, I propose that the stability of moralized beliefs can generate a clearer signal for identifying belief-based allies, while also reassuring potential allies that their shared beliefs are unlikely

to change. The greater efficiency of moralized beliefs in communicating moral beliefs and identifying belief-based allies can be readily observed in the diffusion of moralized content on social media platforms. Specifically, morally laden topics are not only more likely to be shared by individual social media users, they are also more efficiently diffused among networks of users who share the same ideological beliefs (Brady et al., 2017).

Explanatory benefits of a functional account

It is important to note that the purpose of action-based moralization account is not to propose a completely universal pathway to moralization. Indeed, recent research on moralization has increasingly suggested that the phenomenon is complicated and multifaceted, for which the relevant psychological mechanism is strongly influenced by both the context and timeframe (Brandt et al., 2015; Skitka et al., 2017; Wisneski & Skitka, 2016). Instead, the purpose of the action-based moralization account is to fill an explanatory gap that currently exists in terms of how the proximate mechanisms of moralization identified by existing moralization accounts might be activated by factors in an individual's environment.

To this end, the action-based moralization account provides an explanatory framework to bridge the gap that currently exists in the literature between the motivators of moralization, and the psychological and behavioural phenomena associated with issues or actions that have been moralized. More specifically, the central claim of action-based moralization account is that moralization can be motivated by the need to resolve threat-induced anxious-uncertainty, and facilitate future action.

This functional approach to moralization has the potential to contribute to current understanding of subjective moralization by generating novel perspectives for understanding existing proximal accounts. To illustrate, by applying the action-based moralization account, it is possible to theorise that the inconsistent findings around the relevance of disgust and harm to moralization (Brandt et al., 2015; Wisneski et al., 2020; Wisneski & Skitka, 2017)

may be attributable to these mechanisms being manifestations of different moralization functions. Specifically, disgust may be functional to the process of moralization as a signal of threat in the individual's environment (Murray et al., 2019), while the harm-based dyadic template may be more relevant to the ultimate function of efficient communication of moralized beliefs.

The action-based moralization account also provides a functional explanation for the amplifying effect of moral convergence on the outcomes of moralization (Goodwin & Darley, 2012; Mooijman et al., 2018). Specifically, information on moral convergence indicates a clearer potential for group-based coordination, which in turn increases the ultimate functional value of highly stable moralized beliefs.

Novel predictions of the action-based moralization account

The action-based moralization approach provides a powerful explanation for convergent empirical findings showing that, when individuals regard an issue as moral, they not only perceive it as universal (Skitka et al., 2005; van Bavel et al., 2012), and beyond compromise (Skitka, 2010), they also become less tolerant of (Brandt et al., 2016), and actively try to distance themselves from, people who hold alternative viewpoints (Skitka et al., 2005). Importantly, by focusing on the potential function of moralization for the individual, the action-based moralization account is able to generate a number of novel predictions regarding when individuals are likely to moralize, and what the likely outcomes of moralization will be. Specifically the action-based moralization account predicts that: (1) that moralization will be more likely to occur under high levels of threat-appraisal where the individual encounters a state of anxious-uncertainty; and (2) that moralization will result in increased epistemic confidence.

References

Ballone, C., Pacilli, M. G., Teresi, M., Palumbo, R., & Pagliaro, S. (2023). Attitudes moralization and outgroup dehumanization in the dynamic between pro- vs. Anti-vaccines against COVID-19. *Journal of Community & Applied Social Psychology*, *33*(5), 1297–1308. https://doi.org/10.1002/casp.2718

Banisch, S., & Olbrich, E. (2018). An argument communication model of polarization and ideological alignment. *arXiv Preprint arXiv:1809.06134*.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313–7318. https://doi.org/10.1073/pnas.1618923114

Brandt, M. J., Wetherell, G., & Crawford, J. T. (2016). Moralization and Intolerance of Ideological Outgroups. In *The Social Psychology of Morality*. Routledge.

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 U.S. Presidential Election Campaign. *Journal of Social and Political Psychology*, *3*(2), Article 2. https://doi.org/10.5964/jspp.v3i2.434

Cole Wright, J., Cullum, J., & Schwab, N. (2008). The Cognitive and Affective Dimensions of Moral Conviction: Implications for Attitudinal and Behavioral Measures of Interpersonal Tolerance. *Personality and Social Psychology Bulletin*, *34*(11), 1461–1476. https://doi.org/10.1177/0146167208322557

Darwin, C. (1871). *The descent of man: And selection in relation to sex*. John Murray, Albemarle Street.

de Waal, F. B. M., & Waal, F. B. M. de. (1996). *Good Natured*. Harvard University Press. https://doi.org/10.4159/9780674033177

Dechesne, M., & Kruglanski, A. W. (2004). Terror's Epistemic Consequences: Existential
Threat and the Quest for Certainty and Closure. In *Handbook of Experimental
Existential Psychology* (pp. 247–262). Guilford Press.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological
Bulletin*, *139*(2), Article 2. https://doi.org/10.1037/a0029065

Ditto, P. H., & Liu, B. S. (2016). Moral coherence and political conflict. In P. Valdesolo & J.
Graham (Eds.), *Social psychology of political polarization* (pp. 102–122). Routledge.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou,
P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of
misinformation belief and its resistance to correction. *Nature Reviews Psychology*,
1(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Effron, D. A., & Raj, M. (2020). Misinformation and Morality: Encountering Fake-News
Headlines Makes Them Seem Less Unethical to Publish and Share. *Psychological
Science*, *31*(1), 75–87. https://doi.org/10.1177/0956797619887896

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of
moralization: How eating meat becomes a moral issue. *Journal of Personality and
Social Psychology*, *117*(1), Article 1. https://doi.org/10.1037/pspa0000149

FeldmanHall, O., Son, J.-Y., & Heffner, J. (2018). Norms and the Flexibility of Moral
Action. *Personality Neuroscience*, *1*, e15. Cambridge Core.
https://doi.org/10.1017/pen.2018.13

Festinger, L. (1957). *A theory of cognitive dissonance* (pp. xi, 291). Stanford University
Press.

Fiedler, K. (2014). From intrapsychic to ecological theories in social psychology: Outlines of
a functional theory approach. *European Journal of Social Psychology*, *44*(7), Article
7. https://doi.org/10.1002/ejsp.2069

Gampa, A., Wojcik, S. P., Motyl, M., Nosek, B. A., & Ditto, P. H. (2019). (Ideo)Logical

Reasoning: Ideology Impairs Sound Reasoning. *Social Psychological and Personality

Science*, *10*(8), Article 8. https://doi.org/10.1177/1948550619829059

Gershman, S. J. (2019). Uncertainty and Exploration. *Decision (Washington, D.C.)*, *6*(3),

Article 3. https://doi.org/10.1037/dec0000101

Goodwin, G. P. (2015). Moral Character in Person Perception. *Current Directions in

Psychological Science*, *24*(1), 38–44. https://doi.org/10.1177/0963721414550709

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring

objectivism. *Cognition*, *106*(3), Article 3.

https://doi.org/10.1016/j.cognition.2007.06.007

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more

objective than others? *Journal of Experimental Social Psychology*, *48*(1), Article 1.

https://doi.org/10.1016/j.jesp.2011.08.006

Gray, J. A. (1975). *Elements of a two-process theory of learning* (pp. x, 423). Academic

Press.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in

Cognitive Sciences*, *6*(12), Article 12.

Haas, I. J. (2016). The Impact of Uncertainty, Threat, and Political Identity on Support for

Political Compromise. *Basic and Applied Social Psychology*, *38*(3), 137–152.

https://doi.org/10.1080/01973533.2016.1169181

Haas, I. J., & Cunningham, W. A. (2014). The Uncertainty Paradox: Perceived Threat

Moderates the Effect of Uncertainty on Political Tolerance. *Political Psychology*,

35(2), Article 2.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to

moral judgment. *Psychological Review*, *108*(4), Article 4.

https://doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, *316*(5827), 998–1002.

https://doi.org/10.1126/science.1137651

Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An Action-Based Model of

Cognitive-Dissonance Processes. *Current Directions in Psychological Science*, *24*(3),

184–189. https://doi.org/10.1177/0963721414566449

Harmon-Jones, E., & Mills, J. (2019). An introduction to cognitive dissonance theory and an

overview of current perspectives on the theory. In *Cognitive dissonance: Reexamining

a pivotal theory in psychology, 2nd ed* (pp. 3–24). American Psychological

Association. https://doi.org/10.1037/0000135-001

Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., Nash, K., Proulx,

T., & Quirin, M. (2014). Threat and Defense. In *Advances in Experimental Social

Psychology* (Vol. 49, pp. 219–286). Elsevier. https://doi.org/10.1016/B978-0-12-

800052-6.00004-4

Keltner, D., & Gross, J. J. (1999). Functional Accounts of Emotions. *Cognition and Emotion*,

13(5), 467–480. https://doi.org/10.1080/026999399379140

Kovacheff, C., Schwartz, S., Inbar, Y., & Feinberg, M. (2018). The problem with morality:

Impeding progress and increasing divides. *Social Issues and Policy Review*, *12*(1),

Article 1.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational

bases* (pp. xvii, 281). Plenum Press.

Kruglanski, A. W., Orehek, E., Dechesne, M., & Pierro, A. (2010). Lay Epistemic Theory:

The Motivational, Cognitive, and Social Aspects of Knowledge Formation. *Social

and *Personality Psychology Compass*, *4*(10), Article 10.
https://doi.org/10.1111/j.1751-9004.2010.00308.x

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and
"freezing." *Psychological Review*, *103*(2), Article 2. https://doi.org/10.1037/0033-
295X.103.2.263

Liu, B. S., & Ditto, P. H. (2013). What Dilemma? Moral Evaluation Shapes Factual Belief.
Social Psychological and Personality Science, *4*(3), Article 3.
https://doi.org/10.1177/1948550612456045

Louis, W. R., Lizzio-Wilson, M., Cibich, M., McGarty, C., Thomas, E. F., Amiot, C. E.,
Weber, N., Rhee, J., Davies, G., Rach, T., Goh, S., McMaster, Z., Muldoon, O.,
Howe, N., & Moghaddam, F. (2022). Failure Leads Protest Movements to Support
More Radical Tactics. *Social Psychological and Personality Science*, *13*(3), 675–687.
https://doi.org/10.1177/19485506211037296

Luttrell, A., & Togans, L. J. (2021). The Stability of Moralized Attitudes Over Time.
Personality and Social Psychology Bulletin, *47*(4), Article 4.
https://doi.org/10.1177/0146167220935737

McGregor, I., Nash, K., Mann, N., & Phills, C. E. (2010). Anxious uncertainty and reactive
approach motivation (RAM). *Journal of Personality and Social Psychology*, *99*(1),
Article 1. https://doi.org/10.1037/a0019701

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense:
Fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, *28*(3),
Article 3. https://doi.org/10.1016/j.neubiorev.2004.03.005

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative
theory. *Behavioral and Brain Sciences*, *34*(2), 57–74. Cambridge Core.
https://doi.org/10.1017/S0140525X10000968

Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social

    networks and the emergence of violence during protests. *Nature Human Behaviour*,

    2(6), Article 6.

Murray, D. R., Kerry, N., & Gervais, W. M. (2019). On Disease and Deontology: Multiple

    Tests of the Influence of Disease Threat on Moral Vigilance. *Social Psychological*

    and Personality Science, *10*(1), Article 1. https://doi.org/10.1177/1948550617733518

Nelissen, R. M. A., Breugelmans, S. M., & Zeelenberg, M. (2013). Reappraising the Moral

    Nature of Emotions in Decision Making: The Case of Shame and Guilt. *Social and*

    Personality Psychology Compass, *7*(6), 355–365. https://doi.org/10.1111/spc3.12030

Oeberst, A., & Imhoff, R. (2023). Toward Parsimony in Bias Research: A Proposed Common

    Framework of Belief-Consistent Information Processing for a Set of Biases.

    Perspectives on Psychological Science, *18*(6), 1464–1487.

    https://doi.org/10.1177/17456916221148147

Peters, U. (2022). What Is the Function of Confirmation Bias? *Erkenntnis*, *87*(3), 1351–1376.

    https://doi.org/10.1007/s10670-020-00252-1

Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals

    without allies moralize more? *Evolution and Human Behavior*, *34*(2), Article 2.

    https://doi.org/10.1016/j.evolhumbehav.2012.09.006

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A

    review of current definitions, methods, and evidence in moralization research. *Social*

    and Personality Psychology Compass, *13*(12), Article 12.

    https://doi.org/10.1111/spc3.12511

Rozin, P. (1999). The Process of Moralization. *Psychological Science*, *10*(3), Article 3.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping

    between three moral emotions (contempt, anger, disgust) and three moral codes

(community, autonomy, divinity). *Journal of Personality and Social Psychology*,
76(4), Article 4. https://doi.org/10.1037//0022-3514.76.4.574

Ryan, T. J. (2014). Reconsidering Moral Issues in Politics. *The Journal of Politics*, *76*(2),
Article 2. https://doi.org/10.1017/s0022381613001357

Schein, C., & Gray, K. (2016). Moralization and Harmification: The Dyadic Loop Explains
How the Innocuous Becomes Harmful and Wrong. *Psychological Inquiry*, *27*(1),
Article 1. https://doi.org/10.1080/1047840X.2016.1111121

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as Embodied Moral
Judgment. *Personality and Social Psychology Bulletin*, *34*(8), Article 8.
https://doi.org/10.1177/0146167208317771

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary Theory and the
Ultimate–Proximate Distinction in the Human   Behavioral Sciences.
Perspectives on Psychological Science, *6*(1), 38–47.
https://doi.org/10.1177/1745691610393528

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral Conviction: Another Contributor
to Attitude Strength or Something More? *Journal of Personality and Social
Psychology*, *88*(6), 895–917. https://doi.org/10.1037/0022-3514.88.6.895

Skitka, L. J., Hanson, B. E., Morgan, G. S., & Wisneski, D. C. (2021). The Psychology of
Moral Conviction. *Annual Review of Psychology*, *72*(1), 347–366.
https://doi.org/10.1146/annurev-psych-063020-030612

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological
antecedents of threat and challenge appraisal. *Journal of Personality and Social
Psychology*, *73*(1), Article 1.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases.
Science, *185*(4157), 1124–1131.

van Zomeren, M., Postmes, T., & Spears, R. (2012). On conviction's collective

consequences: Integrating moral conviction with the social identity model of

collective action. *British Journal of Social Psychology*, *51*(1), 52–71.

https://doi.org/10.1111/j.2044-8309.2010.02000.x

Wisneski, D. C., Hanson, B. E., & Morgan, G. S. (2020). The roles of disgust and harm

perception in political attitude moralization. *Politics & the Life Sciences*, *39*(2),

Article 2. https://doi.org/10.1017/pls.2020.22

Wisneski, D. C., & Skitka, L. J. (2017). Moralization Through Moral Shock: Exploring

Emotional Antecedents to Moral Conviction. *Personality and Social Psychology

Bulletin*, *43*(2), Article 2. https://doi.org/10.1177/0146167216676479

Yoder, K. J., & Decety, J. (2022). Moral conviction and metacognitive ability shape multiple

stages of information processing during social decision-making. *Cortex*, *151*, 162–

175.

Zaal, M. P., Laar, C. V., Ståhl, T., Ellemers, N., & Derks, B. (2011). By any means

necessary: The effects of regulatory focus and moral conviction on hostile and

benevolent forms of collective action. *British Journal of Social Psychology*, *50*(4),

670–689. https://doi.org/10.1111/j.2044-8309.2011.02069.x

Overview of Empirical Chapters – Testing the Action-Based Moralization Account

The overarching goal of the empirical work conducted in this thesis was to evaluate the initial feasibility of the action-based moralization account. Specifically, the doctoral research program reported in this thesis sought to empirically validate the core premises of the action-based moralization account that: (1) that moralization will be more likely to occur under high levels of threat-appraisal where the individual encounters a state of anxious-uncertainty; and (2) that moralization will result in increased epistemic confidence.

2.2.1. Can moralization be reliably observed as co-occurring with threat-appraisal?

To test the first premise of the action-based moralization account, we developed a multi-method approach designed to evaluate whether moralization could reliably be observed as co-occurring with threat-appraisal. Specifically, we looked at individual differences in sensitivity to threat (Chapter 3), within-person changes in the experience of threat-appraisal emotions (Chapter 4); and between-person differences in threat-appraisals of a specific context (Chapter 5). As a further means of establishing the robustness and generalizability of the findings, each of these studies were also conducted using different contexts. Specifically, Chapter 3 looked at the relationship at an abstract level using highly validated individual differences measures of the threat-sensitivity and moralization constructs. Meanwhile, Chapter 4 applied an ecological momentary assessment method in order to investigate participants' real-world experience of threat-appraisal/moralization dynamics. Finally, Chapter 5 used representative sampling to investigate the threat-appraisal/moralization relationship in the context of a novel real-world event.

2.2.2. The relationship between Moralization and Epistemic Confidence

In Chapter 6, we began to investigate whether moralization does indeed function to increase epistemic confidence. To this end, we used a pre-post design in the context of a real-world political event, to investigate both how existing levels of moralization, and changes in

moralization over time, influenced the level of confidence participants felt about new

information that were potentially threatening to their moralized beliefs.

3

Chapter 3
Who is a Moralizer? Individual Differences in Sensitivity to Threat Predicts Greater Tendency for Moral Condemnation

Chapter 3 Who is a Moralizer? Individual Differences in Sensitivity to Threat Predicts Greater Tendency for Moral Condemnation

3.1. Preface

This chapter presents the results of three studies that investigated the relationship between individual differences in threat-sensitivity and the tendency to moralize third-party actions. The manuscript presented in this chapter was submitted for publication in *Personality and Social Psychology Bulletin* and is presented in author-submitted format.

The first core premise of the action-based moralization account is that the process of moralization can be motivated by threat-appraisal. The broad aim of Chapter 3 was to begin to test this premise using an individual differences approach. Specifically, using two converging measures of individual differences in threat-sensitivity, we sought to investigate whether those who were more likely to appraise situations as threatening, were also more likely to invoke morality in the evaluation of third-party actions across a wide variety of moral domains. While not strictly a test of moralization processes (moralization is by definition a within-person process: see Chapter 2) this individual differences approach allowed us to draw on well-validated measures of threat-sensitivity to test the relevance of the threat-appraisal construct in the context of moral amplification of judgments.

3.2. Manuscript

Declarations

Funding: None.

Conflicts of interest/Competing interests: None.

Availability of data and material:

Code availability:

Ethics approval: This research has been approved by the Human Ethics Research Committee (HREC 1853100.1) at the [BLINDED] and conforms with the Declaration of Helsinki.

Consent to participate: Informed consent was obtained from all individual participants included in the study.

Consent for publication: Consent was obtained from all individual participants included in the study.

Abstract

Across three studies ($N = 665$), we investigated whether trait threat-sensitivity was associated with greater tendency to morally condemn third-party actions using two convergent measures of threat-sensitivity. In Study 1 ($n = 151$) we found that avoidance motivation (BIS) was positively associated, and Sensation-Seeking was negatively associated, with general moral condemnation of third-party actions. We replicated these findings in Study 2 ($n = 206$), and also found that the relationship between BIS/Sensation-Seeking and general moral condemnation remained robust after controlling for trait empathy. In Study 3 ($n = 308$) we found that the relationships between BIS and moral condemnation of third-party actions representing Harm and Sanctity violations, were fully mediated by individuals perceiving these actions as personally threatening (i.e., troubling, worrying and anxiety-inducing). The same was true for the relationship between Sensation-Seeking and moral condemnation of Sanctity (but not Harm) violations.

Keywords: *Moral Judgment, Threat, Individual Differences, Avoidance Motivation, Sensation Seeking*

Individual Differences in Sensitivity to Threat Predicts Greater Tendency for Moral

Condemnation

In today's social-media landscape, it is increasingly common to see individuals publicly

express moral condemnation of the actions and behaviors of unknown others (Brady et al.,

2020). This increasing tendency among members of society to attach moral significance to

others' behaviors has led to heightened awareness of how such *moralization* (see: Rhee,

Schein & Bastian, 2019) can perpetuate and exacerbate polarization (Monroe & Plant, 2019).

Extant research now demonstrates that viewing issues through a moral lens leads individuals

to see their positions as beyond compromise and motivates active distancing from those who

hold contrary positions (Skitka, 2005). Such downstream consequences of moralization mean

that issues viewed through a moral frame are more likely to engender conflict between

opposing sides (Brandt, et al., 2018), and less likely to result in civil debate or compromise

(Delton, et al., 2020).

Third-Party Moral Condemnation

Moral condemnation can emanate from both second and third-person perspectives

(Jones & Nisbett, 1987; Yoder & Decety, 2020). Individuals may express second-person

moral condemnation when they are the victims of a transgression (Kupfer & Giner-Sorolla,

2017). Conversely, third-party moral condemnation is made from the perspective of an

observer. While intuition may suggest that third-party (versus second-party) moral

condemnation may be motivated primarily by concern for the victim (Yoder & Decety, 2020),

or engagement in moral reasoning (see Kohlberg, 1984), a substantial body of evidence

suggests that individuals often engage in third-party moral condemnation for self-oriented

motivations (Miller et al., 2014; Murray et al., 2019; Schnall et al., 2008). For example,

Miller et al. (2014) found that third-party moral condemnation of harm could be best

explained in terms of individuals' aversion to performing the actions themselves, rather than experiencing the consequences of the action.

Disgust-Based Accounts of Moral Condemnation

Although there is substantial evidence that self-oriented motivations may predict third-party moral condemnation, there remains ambiguity regarding the nature of these motivations (see: Skitka et al., 2017; Rhee, et al, 2019). Perhaps the most well-researched self-oriented motivations for third-party condemnation have revolved around the emotion of disgust (e.g. Rozin et al., 2008; Tybur, Lieberman, Kurzban, & DeScioli, 2013).

Disgust-based accounts have taken two broad forms. Pathogen-avoidance accounts (Murry, et al., 2017; Duncan, Schaller, & Park, 2009) suggest that moral condemnation may play a functional role in reducing disease contraction risks by imposing universal standards of behavior upon others. According to this account, individuals may engage in harsher moral condemnation where they feel more vulnerable to pathogen infection (see: Murray, et al., 2019). Other disgust-based accounts posit that a more integral link may exist between the emotion of disgust, and morality (see: Rozin, et al., 2008; Giner-Sorolla & Chapman, 2017; Tybur, et al., 2013). More specifically, these "moral disgust" accounts suggest that, over a period of evolution, moral processes may have co-opted the emotion of disgust due to its general utility as a clear affective signal, making it a distinctly moral emotion (see: Tybur et al., 2013). Such accounts suggest that even the experience of disgust that is incidental to a given judgment may be sufficient to activate harsher moral condemnation (e.g. Schnall et al., 2008).

However, neither of these disgust-based approaches have been able to provide a completely satisfactory account of moral condemnation. While pathogen-based accounts provide a strong functional explanation for increasing moral condemnation (see: Murray et al., 2019), they are unable to explain why increased moral condemnation can also occur in the

absence of pathogen cues (e.g. Landy & Piazza, 2019). On the other hand, while moral-disgust accounts provide a theoretical explanation that may be applicable to a broad range of moral phenomena, recent studies have challenged the notion that moral condemnation is reducible to the single emotion of disgust (Cameron et al., 2015; Landy & Goodwin, 2015). For example, a meta-analysis conducted by Landy and Goodwin (2015) found only a small effect of disgust inductions on increased moral condemnation, suggesting that the central role of disgust may have been over-stated. Furthermore, in direct contradiction of the specific role of disgust in moral condemnation, Landy and Piazza (2019) found that general sensitivity to a range of negative affective states (not simply to disgust) was associated with greater moral condemnation. While these studies do not unequivocally deny any role of disgust, they provide evidence that disgust may not be the primary elicitor of moral condemnation. Such findings thus call for the identification of further mechanisms underlying moral condemnation. In the present studies, we start with a mechanism distal to disgust—namely, the general motivation to avoid threats in one's environment—may underly moral condemnation.

General Threat Sensitivity

The detection of threat cues in the environment triggers a series of affective, cognitive, behavioral and physiological responses resulting in heightened arousal and preparedness to respond to danger (Denefrio & Dennis-Tiwary, 2018; Tomaka et al., 1997). While this threat-response mechanism is universal, evidence suggests that individuals can differ in their degree of sensitivity to threat cues, and in their motivation for threat avoidance (Cools et al., 2005; Denefrio & Dennis-Tiwary, 2018; Gray, 1990; Grillon et al., 2017; Kramer et al., 2019; Lissek & Powers, 2003; Yancey et al., 2015). The most influential conceptualization of such individual differences in threat-sensitivity has been Gray's Reinforcement Sensitivity Theory of Personality (RST: Gray, 1978, McNaughton & Gray,

2001), which delineates two fundamental axes of motivation along which individuals may differ—namely their sensitivity to appetitive cues (behavioral activation system: BAS), and aversive cues (behavioral inhibition system: BIS). According to RST, the BIS network is involved in the regulation of motives to avoid fearful, novel or unpleasant experiences or stimuli, as well as physical and emotional responses to aversive cues (Carver & White, 1994; Gray, 1990). Consistent with the notion that BIS represents a fundamental dimension of threat-processing, individual difference measures of BIS (or avoidance motivation: Carver & White, 1994) have been found to co-occur with biological and reflexive forms of threat response such as enhanced neural processing of fearful faces (Cools et al., 2005), increased startle reflex responsiveness when presented with threatening stimuli (Caseras et al., 2006; Lissek & Powers, 2003; Yancey et al., 2015), and higher levels of baseline cortisol (Tops & Boksem, 2011).

Threat Sensitivity and Moral Condemnation

General sensitivity to threat has long been discussed as relevant to moral processing (Laham & Corless, 2016; Wright & Baril, 2013). Chief among these has been the Model of Moral Motives (MMM: Janoff-Bulman & Carnes, 2009), which suggests that moral regulation may be divided broadly into prescriptive and proscriptive forms. According to MMM, proscriptive morality acts to regulate failures to inhibit actions which may lead to negative outcomes (Janoff-Bulman & Carnes, 2009). By extension, MMM suggests that individuals who are particularly motivated by threat avoidance, would be more likely to engage in proscriptive ("condemning past wrongs") over prescriptive ("encouraging future actions") moral regulation (Janoff-Bulman & Carnes, 2009).

When applied to the context of moral condemnation, MMM should predict that individuals higher in threat-sensitivity, would be more likely to morally condemn third-party actions as a means of regulating behavior that they find aversive. However, to date, little

work has investigated whether this is the case, and the proximal mechanisms by which this might occur. Further, while MMM accounts have suggested that mechanisms of proscriptive moral regulations apply predominantly to social conservatives (Janoff-Bulman & Carnes, 2013; Janoff-Bulman et al., 2009), more recent evidence (Bakker et al., 2020; Crawford, 2017) has questioned past assertions that social conservatives (vs liberals) are defined by higher threat-sensitivity (Thórisdóttir & Jost, 2011; Wakslak et al., 2007).

Overview of Current Research

In the present studies, we investigated whether individual differences in basic threat-sensitivity were related to a tendency for moral condemnation, controlling for social conservativism. We measured basic threat-sensitivity using Carver & White (1994)'s well-established measure of BIS. Further, in order to establish convergent validity around the construct of basic threat-sensitivity as the relevant mechanism, we additionally used Zuckerman (1984)'s construct of Sensation Seeking. While not originally developed as a specific measure of threat-sensitivity, individual differences in Sensation Seeking have been found to moderate basic threat-responding such as startle-reflex responses to threatening stimuli (Kramer et al., 2019; Kramer et al., 2012; Lissek & Powers, 2003). We therefore predicted that both individuals high in BIS and individuals low in Sensation-Seeking would be more likely to engage in general moral condemnation of third-party actions (Studies 1 & 2). Throughout these studies, we also examined the applicability of this proposed mechanism across multiple moral domains, using a validated and standardized measure (Clifford et al. 2015) of sensitivity to violations of each domain originating from moral foundations theory (Graham et al. 2009). Finally, in Study 3, we sought to more directly investigate whether perceiving third-party actions as personally threatening would mediate the relationship between trait threat-sensitivity and moral condemnation.

Study 1

In Study 1 we sought to investigate the relationship between moral condemnation of third-party actions, and trait threat-sensitivity. More specifically, we looked at whether two convergent measures of threat-sensitivity (BIS and Sensation-Seeking) were associated with moral condemnation of third-party actions which violated a range of moral foundations domains.

Method

Participants

A convenience sample of one-hundred and fifty-one participants (Male = 75, Female = 76) aged between 20 and 71 years ($M = 33.95$, $SD = 10.28$) were recruited via Amazon's Mechanical Turk platform (Buhrmester et al., 2011). Power analyses conducted using G*Power (Faul, Erdfelder, Lang, Buchner, 2007), revealed that this sample was sufficient to detect a medium correlation ($r = .30$) with 90% power.

Materials and Procedure

Social Conservativism. Social conservativism was measured using a one-item question: "Please indicate your political beliefs from left/liberal to right/conservative on social issues (e.g., immigration, homosexual marriage, abortion)". Responses were indicated on a 7-point scale (1 = "*Left/liberal*" to 7 = "*Right/conservative*"), such that high scores indicated more social conservativism

Behavioral Inhibition Sensitivity. Trait avoidance motivation was assessed using the seven (7) item BIS subscale from the Behavioral Inhibition (BIS)/Behavioral Activation Scales (BAS) (Carver & White, 1994). Participants responded to the items on a scale from 1 (*Very true for me*) to 4 (*Very false for me*). Two items on the were reverse-scored (e.g. "I have very few fears compared to my friends"). Participants' trait BIS was calculated as their average ratings across items such that higher scores indicated higher trait BIS ($\alpha = .80$).

Sensation-Seeking. Trait Sensation-Seeking was measured using the 40-item Sensation-Seeking Scale V (Zuckerman, 1984). Each item in the Sensation-Seeking Scale presents participants with two choices, and asks them to select the option that most describes their likes or feelings (e.g. "I prefer friends who are excitingly unpredictable" vs. "I prefer friends who are reliable and predictable"). Participants' responses were coded such that those who selected the high Sensation-Seeking option for a given item were scored as "1" and those who selected the alternative were scored as "0". Participants' overall level of Sensation-Seeking was calculated as the sum of all item scores, such that higher scores indicated higher trait Sensation-Seeking ($\alpha = .82$).

Moral Condemnation. Participants' trait propensity for moral condemnation was measured using the Moral Foundations Vignettes (MFV: Clifford, Iyengar, Cabeza, and Sinnott-Armstrong, 2015). The MFV are a set of ninety (90) vignettes portraying intentional actions performed by a third-party, which represent violations of each of the domains from moral foundations theory (see: Iyer, Koleva, Graham, Ditto & Haidt, 2012) – specifically: *Harm* (27 items: e.g. "you see a teenage boy chuckling at an amputee he passes by while on the subway"); *Fairness* (12 items: e.g. "you see a student copying a classmate's answer sheet on a final exam"), *Loyalty* (14 items: e.g. "you see an employee joking with competitors about how badly his company did last year"), Liberty (11 items: e.g. "you see a man blocking his wife from leaving home or interacting with others"), Authority (14 items: e.g. "you see a woman refusing to stand when the judge walks into the courtroom"), and Sanctity (12 items: e.g. "you see a teenager urinating in the wave pool at a crowded amusement park").

Participants were asked to indicate how morally wrong they perceive each action to be on a 5-point scale ranging from 1 (not at all morally wrong) to 5 (extremely morally wrong). As has been done previously (see: Murray, et al., 2017), we created a moral condemnation composite, by taking the average across all domains such that high scores

indicated greater tendency for general moral condemnation (α = .98). Participants' responses were also averaged within each domain, such that higher scores indicated greater moral condemnation of actions violating each domain – i.e., Harm (α = .95), Fairness (α = .92), Liberty (α = .93), Loyalty (α = .96), Authority (α = .96), and Sanctity (α = .93).

Results

Zero-order correlations between measures of threat-sensitivity, social conservativism and moral condemnation across the six moral foundations are shown in Table 1. Consistent with general threat-sensitivity account, we found that both Sensation Seeking (r = -.43, p < .001) and BIS (r = .22, p = .006) were associated with general moral condemnation. These relationships remained significant after controlling for social conservativism (see Table 1), which itself was positively correlated with general moral condemnation, r = .25, p = .002.

Looking at moral condemnation in each specific domain, as predicted by moral foundations theory, social conservativism was positively associated with greater moral condemnation of actions violating the "binding foundations" (see: Haidt & Graham, 2007) of Authority (r = .33, p < .001), Loyalty (r = .36, p <.001) and Sanctity (r = .30, p < .001). We also found that trait Sensation Seeking was negatively associated with moral condemnation of actions violating the Harm, Farness, Liberty, Authority, Loyalty and Sanctity domains ($|rs|$>.18, ps < .032), while BIS was positively associated with moral condemnation of the Harm, Fairness, Liberty and Sanctity domains, $|rs|$>.18, ps < .024. Again, we found that these relationships remained after controlling for social conservativism (see Table 1).

Discussion

In Study 1, we found initial evidence that two convergent measures of threat sensitivity (BIS and Sensation Seeking) were associated with general moral condemnation, even after controlling for political conservativism. We found that Sensation-Seeking was more strongly related to general moral condemnation than BIS. Upon examination of moral

condemnation at the domain-level we found that trait BIS was more weakly associated with moral condemnation across all domains when compared to Sensation-Seeking, and further that BIS did not relate to moral condemnation of Loyalty and Authority violations.

Study 2

In Study 2, we sought to replicate the findings in Study 1 with a larger sample. We also sought to extend these findings by accounting for the potential influence of trait differences in other-oriented concern for the victim. This was particularly important as empathy has previously been demonstrated to predict moral condemnation (Yoder & Decety, 2020). Whilst not discounting a potential role of empathy, we predicted that the relationship between self-oriented threat-sensitivity and moral condemnation would remain significant after controlling for trait empathy.

Method

Participants

A convenience sample of two-hundred and six participants (114 male, 92 female) aged between 18 and 71 years ($M = 32.9$, $SD = 9.2$) were recruited via Amazon's Mechanical Turk platform (Buhrmester et al., 2011). Power analyses conducted using G*Power (Faul et al., 2007) revealed that this sample gave us sufficient power to detect, the smallest effect of interest on general moral condemnation in Study 1 ($r = .22$) with 90% power.

Materials and Procedure

Sensation-Seeking ($\alpha = .80$), BIS ($\alpha = .87$) and Moral Condemnation ($\alpha = .98$) were measured using the corresponding scales described in Study 1.

Perspective Taking and Empathic Concern. Trait empathy was measured using the Trait Perspective Taking and Empathic Concern subscales from the Interpersonal Reactivity Index (IRI: Davis, 1980; Keaton, 2017). A total of 14 items from the IRI were administered, with 7-items measuring perspective taking (e.g. "I sometimes try to understand my friends

better by imagining how things look from their perspective"), and 7 items measuring empathic concern (e.g. "I often have tender, concerned feelings for people less fortunate than me."). Two reverse-scored items were included in both the perspective taking (e.g. "I sometimes find it difficult to see things from the "other guy's" point of view."), and empathic concern scales (e.g. "Other people's misfortunes do not usually disturb me a great deal."). Participants responded to the items on a scale from 1 (*Does not describe me well*) to 5 (*Describes me very well*). Responses were averaged such that high scores indicated either high trait perspective-taking ($\alpha = .89$) or high trait empathic concern ($\alpha = .87$).

Results

Zero-order correlations between the Threat-Sensitivity variables and Moral Condemnation are presented in Table 2. Replicating the results of Study 1: social conservativism was positively associated with moral condemnation of Authority, Loyalty and Sanctity violations; trait Sensation Seeking was negatively associated with moral condemnation of all six moral foundations vignettes ($|rs|>.17$, $ps < .037$); and, trait BIS was positively associated with the same moral foundations as Study 1 ($|rs|>.18$, $ps < .010$). In addition, trait BIS was also positively associated with moral condemnation of vignettes representing violations of the Authority domain ($r = 16$, $p = .022$).

We conducted hierarchical multiple regression analyses to investigate the relationship between trait threat-sensitivity and moral condemnation after controlling for Perspective Taking and Empathic Concern. Separate regression analyses were performed for Sensation-Seeking and BIS as predictors of general moral condemnation, as well as moral condemnation in each domain. For each model, the threat-sensitivity variables were entered at Block 1, and Empathic Concern, Perspective Taking and Social Conservativism were entered at Block 2.

First, examining the models predicting general moral condemnation, both Sensation Seeking ($\beta = -.19$, $p = .003$) and BIS ($\beta = .28$, $p < .001$), remained significant predictors of general moral condemnation after controlling for Social Conservativism, Empathic Concern, and Perspective Taking (see Tables 3 & 4).

Examining each specific domain, trait Sensation Seeking remained a significant predictor for the Harm ($\beta = -.20$, $p = .003$), Fairness ($\beta = -.17$, $p = .043$), and Sanctity domains ($\beta = -.22$, $p = .001$), however, was no longer a significant negative predictor of the Authority ($\beta = -.10$, $p = .143$) and Loyalty ($\beta = -.10$, $p = .136$) domains, after controlling for social conservativism, Empathic Concern, and Perspective Taking (see Table 3). On the other hand, the pattern of results remained the same for trait BIS after controlling for social conservativism, Empathic Concern, and Perspective Taking. Specifically, even after adding the control variables, trait BIS remained a positive significant positive predictor of Harm ($\beta = .36$, $p < .001$), Fairness ($\beta = .21$, $p = .003$), Liberty ($\beta = .19$, $p = .007$), Authority ($\beta = .20$, $p = .004$) and Sanctity ($\beta = .24$, $p = .001$) violations (see Table 4).

Discussion

In Study 2, we were able to replicate findings in Study 1 that individual differences in threat-sensitivity are associated with general moral condemnation. Specifically, we found that general moral condemnation was negatively associated with trait Sensation Seeking and positively associated with trait BIS. These general relationships remained significant even after controlling for political conservativism, and individual differences in other-oriented concerns measured by Empathic Concern and Perspective Taking.

At the level of individual moral foundation domains, the pattern of associations between trait threat-sensitivity and moral condemnation were again similar to Study 1. However, there was some divergence in the pattern of results between BIS and Sensation Seeking, after controlling for trait Empathic Concern and Perspective Taking. Specifically,

we also found that the relationship between trait BIS and moral condemnation remained significant across all moral foundation domains even after controlling for individual differences in Perspective Taking and Empathic Concern, suggesting that the BIS/moral-condemnation relationship may be driven by more self-oriented responses to the relevant third-party actions. On the other hand, the Sensation-Seeking/moral-condemnation relationships became non-significant for Loyalty and Authority domains after controlling for trait empathic concern and perspective taking.

Given that for both Studies 1 & 2, the relationship between BIS and moral condemnation of Authority and Loyalty violations were weaker and more inconsistent than for the other domains, this divergence in the pattern of results between BIS and Sensation-Seeking for moral condemnation of these two domains, may suggest that these relationships are driven by separate mechanisms. Specifically, the results of Study 2 suggest that the relationship between Sensation Seeking and moral condemnation of Loyalty and Authority violations may be accounted for by differing levels of Empathic Concern.

Study 3

In Study 3, we sought to gain greater insight into the mechanisms underlying the relationship between threat-sensitivity and moral condemnation. Specifically, we sought to investigate whether the relationships between threat-sensitivity and moral condemnation observed in Studies 1 & 2 were indeed driven by a greater feeling of self-oriented threat upon reading the vignettes. Further, we sought to evaluate the importance of a self-oriented threat mechanism, in relation to mechanisms proposed by current influential theories of moralization. In particular, we investigated the applicability of disgust-based accounts (Pizarro et al., 2011; Schnall et al., 2008), and the influential harm-based dyadic account of moralization (Schein & Gray, 2015). The harm-based dyadic account proposes that perceptions of harmfulness and immorality form a dyad of mutual reinforcers. Thus, when

something is perceived to be harmful, it comes to also be seen as immoral, which in turn increases how harmful it is perceived to be, and so on (Schein & Gray, 2015). According to this account, the proximal mechanism for the relationship between threat-sensitivity and moral condemnation may be the perception of the action as potentially harmful to others, rather than a self-oriented appraisal that the act is threatening.

In order to test the utility of a self-oriented general threat-sensitivity account, beyond these existing explanations, we focused on a narrower selection of moral condemnation vignettes drawn from those used in Studies 1 & 2. Specifically, we selected vignettes representing violations of the Harm and Sanctity domains, as these were the most relevant to the aforementioned harm-based and disgust-based accounts of moralization. We then sought to evaluate the applicability of each of these theories in mediating the relationship between threat-sensitivity and moral condemnation, by prompting participants to answer questions gauging each relevant mechanism prior to indicating the moral wrongness of the third-party action portrayed in each vignette. We predicted that both BIS and Sensation Seeking would have an indirect effect on moral condemnation *via* feelings of self-oriented threat for both Harm and Sanctity violations. As with Study 2, we did not predict that a self-oriented threat-based mechanism would supersede harm-based or disgust-based mechanisms, but rather that it would remain an important mediator, even after controlling for them. We further predicted that perceptions of harmfulness would be a more important mediator for Harm violations than for Sanctity violations, while the inverse would be true for the mediating effect of feelings of disgust.

Method

Participants

A convenience sample of three-hundred and two participants (Male = 179, Female = 123) aged between 18 and 74 years ($M = 23.03$, $SD = 11.28$) were recruited via Amazon's

Mechanical Turk platform (Buhrmester et al., 2011). Power analyses conducted using G*Power (Faul et al., 2007) revealed that this sample gave us sufficient power to detect, the smallest relationship of interest from Study 2 (between BIS and Sanctity Condemnation: r =.18) with 90% power.

Materials and Procedure

Sensation-Seeking and BIS were measured using the corresponding scales described in Study 1. After completing the threat-sensitivity pre-measures, each participant was presented with a random subset of four Harm (or Sanctity) violation vignettes and asked to rate each vignette's moral wrongness. Thereafter, each vignette was presented on a separate page of survey in randomized order, and participants were asked to indicate their experience of our proposed mediating variables, in relation to the vignette. The same procedure was repeated with a random subset of four Sanctity (or Harm) violation vignettes. The presentation order of Harm versus Sanctity violations was randomized between participants. Separate moral condemnation scores were calculated for judgments of actions violating the Harm and Sanctity domains. Specifically, for each domain, moral condemnation was calculated as the average moral wrongness rating given by each participant across the four vignettes that they were presented with.

Proposed Mediators

Participants responded to all mediator questions on a 7-point scale (1 = "*Strongly Disagree*", 7 = "*Strongly Agree*").

Self-Oriented Threat-Appraisals. Participants' self-oriented threat-appraisals in response to the vignettes were assessed using the threat-appraisal emotions used by (Michel et al., 2016). Specifically, participants were asked to indicate the extent to which they personally found the behaviors "Worrying", "Anxiety-Inducing" and "Troublesome".

Responses on these items were averaged such that high scores indicated higher feelings of self-oriented threat after viewing the Harm (α = .84), and Sanctity (α = .87) violations.

Perceived Harm. Participants' perceptions of the harmfulness of each vignette was assessed with a single item: "To what extent do you think this behavior may be harmful to others?". High scores on this item indicated a greater perception that the action may be harmful to others.

Disgust. Participants' feelings of disgust at each vignette was assessed with a single item: "To what extent do you think this behavior is disgusting?", such that higher scores indicated greater feelings of disgust at the action.

Results

Zero-order correlations between the Threat-Sensitivity variables, proposed mediators and Moral Condemnation are presented in Table 5.

To investigate the relationships between Threat-Sensitivity variables, proposed mediators and moral condemnation, linear mixed-effects model analyses were performed in *R* using the *Lme4* package (Bates et al., 2015), and significance values were obtained using the *LmerTest* package (Kuznetsova et al., 2017). For the purpose of these analyses, all variables were mean-centered and standardized. Participants and presented vignettes were included as random intercepts in order to account for participants viewing different sets of three (3) Harm, and three (3) Sanctity violation vignettes. Replicating Studies 1 & 2, we found that, controlling for political conservativism, BIS was a significant positive predictor of moral condemnation of actions representing both Harm (b = .21, SE = .06, p < .001) and Sanctity violations, b = .14, SE = .06, p = .009. Similarly, controlling for political orientation, Sensation Seeking was a significant negative predictor of moral condemnation of actions representing both Harm (b = -14, SE = .06, p = .016) and Sanctity violations, b = -.21, SE = .05, p < .001.

We then investigated whether individual differences in BIS and Sensation Seeking predicted participants' experience of each of the theoretical mediators. As hypothesized, we found that BIS was a significant positive predictor of participants' feeling threat-appraisal emotions upon reading both the Harm ($b = .13$, $SE = .04$, $p < .001$) and Sanctity vignettes, $b = .25$, $SE = .05$, $p < .001$. Similarly, Sensation Seeking was a significant negative predictor of participants' feeling threat-appraisal emotions upon reading both the Harm ($b = -.11$, $SE = .06$, $p = .04$) and Sanctity vignettes, $b = -.23$, $SE = .05$, $p < .001$.

Hayes' PROCESS macro for SPSS (Hayes, 2013; Model 4) was used to investigate the indirect effects of trait threat-sensitivity on moral condemnation via self-oriented threat. Four separate model analyses were performed for BIS and Sensation-Seeking, predicting moral condemnation of Harm and Sanctity domains. In all models, social conservativism was entered as a covariate. As predicted, results from 95% confidence intervals (with 5000 resamples) revealed significant indirect effects of BIS on moral condemnation of both Harm and Sanctity violations via experience of self-oriented threat (see *Figure 1*). Similarly, there was a significant indirect effect of Sensation-Seeking on moral condemnation of Sanctity violations via self-oriented threat, however this indirect effect was not significant for moral condemnation of Harm violations. Notably, after controlling for social conservativism, there was a significant total effect of BIS on Sanctity violations (see *Figure 1*), suggesting that social conservativism may have been suppressing this relationship at the zero-order level.

Next, to evaluate the utility of a self-oriented threat account over and above harm-based and disgust-based accounts, we conducted four separate parallel mediation model analyses (Hayes, 2013; Model 4), representing the indirect effects of BIS and Sensation-Seeking on Harm and Sanctity violations. In each model, all three proposed mediators were entered simultaneously to investigate the relative importance of each indirect effect. For the purposes of these analyses, social conservativism was entered as a covariate. Results from

95% confidence intervals (with 5000 resamples) revealed significant indirect effects of BIS on moral condemnation of Harm violations via self-oriented threat, perceived harm, and feelings of disgust (see Table 6). The indirect effects of both Sensation-Seeking and BIS on moral condemnation of Sanctity violations were only significant *via* self-oriented threat and disgust, and not *via* perceived harm. No significant indirect effects were observed for the relationship between Sensation-Seeking and moral condemnation of Harm (see Table 6).

General Discussion

In this research, we aimed to investigate the potential relationship between individuals' threat-sensitivity, and their tendency for moral condemnation. In Studies 1 and 2, we predicted that trait threat-sensitivity both in terms of avoidance motivation (measured by BIS), and Sensation-Seeking would be associated with third-party moral condemnation. As predicted, we found general moral condemnation of third-party actions were associated positively with BIS and negatively with Sensation-Seeking. In Study 2, we also found that these relationships remained significant after controlling for trait empathy – suggesting that this relationship may be self-oriented, rather than emanating from concern for the victim. This pattern of results was largely reflected at the level of individual moral domains, with the exception of relationships between Sensation-Seeking and moral condemnation of Loyalty and Authority violations, which became non-significant after controlling for trait empathy.

Finally, in Study 3, we predicted that the relationships between threat-sensitivity and moral condemnation would be mediated by self-oriented threat. As predicted, we found that the relationship between BIS and moral condemnation of violations in both Harm and Sanctity domains were fully mediated by perceptions that the actions were personally threatening. Importantly, this indirect effect via self-oriented threat, remained significant in a model that accounted for self-reported disgust at the actions, and how harmful the actions were perceived to be. However, while the indirect effect of Sensation-Seeking via self-

oriented threat was significant for condemnation of Sanctity-violations, we did not observe this for Harm violations, and further found no indirect effects *via* harm-perception or feelings of disgust.

Threat-Sensitivity as a Broader Account of Moral Condemnation

This research makes important contributions to understanding regarding moral condemnation. First, it demonstrates the potential role of individual differences in threat-sensitivity – a basic motivational construct – in explaining tendencies for moral condemnation. Notably, the current research provides convergent evidence of the relationship between trait threat-sensitivity and moral condemnation, using two theoretically distinct measures; that is, BIS and Sensation Seeking.

Since its development, trait BIS has been found to relate to a range of pathologies such as anxiety and depression (Bijttebier et al., 2009). Further, Carver & White (1994)'s trait BIS measure has demonstrated remarkable utility in predicting variance between individuals in terms of instinctive behaviors such as startle-eyeblink responses when cued to expect negative or threatening stimuli (Caseras et al., 2006; Crowell et al., 2015); and even neurochemical processes such as increasing the effect of serotonin depletion on amygdala activation in response to fearful stimuli (Cools et al., 2005). While (unlike BIS) Sensation Seeking has not traditionally focused specifically on aversion to threats, a large body of research has demonstrated that trait Sensation Seeking can similarly influence basic threat-response processes such as startle eye-blink responses (Lissek et al., 2005; Lissek & Powers, 2003), and the magnitude of event-related potential (ERP) responses to threatening stimuli (Zheng et al., 2011).

By showing that individual differences in both BIS and Sensation-Seeking are associated with moral condemnation across a wide range of moral foundations, our study highlights the potential application of the wealth of research on threat-sensitivity and

motivation to understanding moral processes. It also corroborates recent research questioning the centrality of disgust as a mechanism for third-party moral condemnation (e.g. Landy & Goodwin, 2015; Landy & Piazza, 2019). Furthermore, by demonstrating that self-oriented threat appraisals remain an important mechanism for moral condemnation in both Harm and Sanctity domains – even after controlling for perceptions of harm, and feelings of disgust – we demonstrate that this threat-sensitivity mechanism can provide insights that go beyond current influential accounts of moralization (e.g. Schein & Gray, 2017; Tybur, et al., 2013).

Finally, as far as we are aware, this is the first study to demonstrate a consistent link between threat-sensitivity and moral condemnation while controlling for social conservativism. Our findings thus go beyond existing threat-based accounts of morality, which apply this construct to account for differences between conservatives and liberals (Janoff-Bulman & Carnes, 2013; Thórisdóttir & Jost, 2011). Our more general threat-based account brings theorizing in the moral domain in line with more recent evidence challenging the notion that conservatives are more sensitive to threat (Bakker et al., 2020; Crawford, 2017)

Limitations and Future Directions

While there was a consistent relationship between BIS, Sensation-Seeking, and general moral condemnation, some inconsistencies were observed at the level of individual moral domains. Specifically, the relationship between Sensation-Seeking and moral condemnation of Authority and Loyalty violating actions became non-significant after the controlling for empathy. Such inconsistencies may be attributable to Sensation-Seeking being a less direct measure of threat-sensitivity when compared to BIS. While low Sensation-Seeking has been linked to implicit threat-avoidance behaviors, the construct was initially conceptualized as the tendency to be excited by novel experiences and willingness to take risks (Zuckerman, 1984). It is possible that negative associations between Sensation-Seeking

and moral condemnation of Authority and Loyalty violations were driven by the desire for novelty leading to less aversion for actions undermining group hierarchies (see: Graham et al., 2009). Future studies could investigate this by applying a factor analytical approach, wherein items loading onto a threat-sensitivity factor could be drawn from existing measures (see: Kramer et al., 2012), and evaluated against items loading onto novelty-seeking, in terms of their relationships with moral condemnation.

A further limitation of the current study was the use of a correlational design, which – although effective in investigating individual differences – did not allow for inferences regarding dynamic processes. Thus, we provide initial evidence for why individuals may differ in their tendency to morally condemn third-party actions, our findings do not provide insights into contextual factors that may lead individuals to moralize the actions of others. However, such insights carry important implications for public policy, especially in the context of societies with high levels of affective polarization (Brady et al., 2020), and thus is an important direction for future work.

References

Bakker, B. N., Schumacher, G., Gothreau, C., & Arceneaux, K. (2020). Conservatives and

   liberals have similar physiological responses to threats. *Nature Human Behaviour*, 1-

   9.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models

   Using lme4." *Journal of Statistical Software*, **67**(1), 1–48. doi: 10.18637/jss.v067.i01.

Bijttebier, P., Beck, I., Claes, L., & Vandereycken, W. (2009). Gray's Reinforcement

   Sensitivity Theory as a framework for research on personality–psychopathology

   associations. *Clinical Psychology Review, 29*(5), 421-430.

Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD Model of Moral

   Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized

   Content Online. *Perspectives on Psychological Science, 15*(4), 978-1010.

   https://doi.org/10.1177/1745691620917336

Brandt, M. J., Crawford, J. T., & Van Tongeren, D. R. (2019). Worldview conflict in daily

   life. *Social Psychological and Personality Science*, *10*(1), 35-43.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new

   source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science,*

   6(1), 3-5.

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A Constructionist Review of Morality

   and Emotions:No Evidence for Specific Links Between Moral Content and Discrete

   Emotions. *Personality and Social Psychology Review, 19*(4), 371-394.

   https://doi.org/10.1177/1088868314566683

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and

   affective responses to impending reward and punishment: The BIS/BAS Scales.

Journal of Personality and Social Psychology, 67(2), 319-333.

https://doi.org/10.1037/0022-3514.67.2.319

Caseras, F., Fullana, M., Riba, J., Barbanoj, M., Aluja, A., & Torrubia, R. (2006). Influence of

individual differences in the Behavioral Inhibition System and stimulus content (fear

versus blood-disgust) on affective startle reflex modulation. *Biological psychology,*

72(3), 251-256.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015, December 01). Moral

foundations vignettes: a standardized stimulus database of scenarios based on moral

foundations theory [journal article]. *Behavior research methods, 47*(4), 1178-1198.

https://doi.org/10.3758/s13428-014-0551-2

Cools, R., Calder, A. J., Lawrence, A. D., Clark, L., Bullmore, E., & Robbins, T. W. (2005).

Individual differences in threat sensitivity predict serotonergic modulation of

amygdala response to fearful faces. *Psychopharmacology, 180*(4), 670-679.

Crawford, J. T. (2017). Are conservatives more sensitive to threat than liberals? It depends on

how we define threat and conservatism. *Social cognition, 35*(4), 354-373.

Crowell, A., Page-Gould, E., & Schmeichel, B. J. (2015, Apr). Self-affirmation breaks the

link between the behavioral inhibition system and the threat-potentiated startle

response. *Emotion, 15*(2), 146-150.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS*

Catalog of Selected Documents in Psychology, 10, 1–19.

Delton, A. W., DeScioli, P., & Ryan, T. J. (2020). Moral obstinacy in political

negotiations. *Political Psychology*, *41*(1), 3-20.

Denefrio, S., & Dennis-Tiwary, T. A. (2018). Threat Sensitivity. *Encyclopedia of Personality*

and Individual Differences. Springer, Cham.

Duncan, L. A., Schaller, M., & Park, J. H. (2009). Perceived vulnerability to disease:
Development and validation of a 15-item self-report instrument. *Personality and
Individual differences*, *47*(6), 541-546.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets
of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.

Gray, J. A. (1990). Brain Systems that Mediate both Emotion and Cognition. *Cognition and
Emotion, 4*(3), 269.

Grillon, C., Robinson, O. J., Krimsky, M., O'Connell, K., Alvarez, G., & Ernst, M. (2017,
Mar). Anxiety-mediated facilitation of behavioral inhibition: Threat processing and
defensive reactivity during a go/no-go task. *Emotion, 17*(2), 259-266.

Giner-Sorolla, R., & Chapman, H. A. (2017). Beyond purity: Moral disgust toward bad
character. *Psychological science*, *28*(1), 80-91.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A. (2007). G*Power 3: A flexible statistical
power analysis program for the social, behavioral, and biomedical sciences. Behavior
Research Methods, 39, 175–191Janoff-Bulman, R., & Carnes, N. C. (2013).
Personality and Social Psychology Review, 17(3), 219.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality:
Two faces of moral regulation. *Journal of Personality and Social Psychology, 96*(3),
521.

Jones, E. E., & Nisbett, R. E. (1987). The actor and the observer: Divergent perceptions of
the causes of behavior. Preparation of this paper grew out of a workshop on
attribution theory held at University of California, Los Angeles, Aug 1969.

Keaton, S.A. (2017). Interpersonal Reactivity Index (IRI). In D.L. Worthington and G.D.
Bodie (Eds.), *The Sourcebook of Listening Research.* (pp. 340-347).

Kramer, M. D., Patrick, C. J., Hettema, J. M., Moore, A. A., Sawyers, C. K., & Yancey, J. R. (2019, 2020/04/01). Quantifying Dispositional Fear as Threat Sensitivity: Development and Initial Validation of a Model-Based Scale Measure. *Assessment, 27*(3), 533-546. https://doi.org/10.1177/1073191119837613

Kramer, M. D., Patrick, C. J., Krueger, R. F., & Gasperi, M. (2012). Delineating physiologic defensive reactivity in the domain of self-report: phenotypic and etiologic structure of dispositional fear. *Psychological Medicine, 42*(6), 1305-1320. https://doi.org/10.1017/S0033291711002194

Kupfer, T. R., & Giner-Sorolla, R. (2017). Communicating moral motives: The social signaling function of disgust. *Social Psychological and Personality Science, 8*(6), 632-640.

Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software, 82*(13), 1–26. doi: 10.18637/jss.v082.i13.

Laham, S. M., & Corless, C. (2016). Threat, morality and politics: A differentiated threat account of moral and political values.

Landy, J. F., & Goodwin, G. P. (2015, 2015/07/01). Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence. *Perspectives on Psychological Science, 10*(4), 518-536. https://doi.org/10.1177/1745691615583128

Landy, J. F., & Piazza, J. (2019). Reevaluating Moral Disgust: Sensitivity to Many Affective States Predicts Extremity in Many Evaluative Judgments. *Social Psychological and Personality Science, 10*(2), 211-219. https://doi.org/10.1177/1948550617736110

Lissek, S., Baas, J. M., Pine, D. S., Orme, K., Dvir, S., Rosenberger, E., & Grillon, C. (2005). Sensation Seeking and the Aversive Motivational System. *Emotion, 5*(4), 396.

Lissek, S., & Powers, A. S. (2003). Sensation seeking and startle modulation by physically threatening images. *Biological psychology, 63*(2), 179-197.

Michel, A., Turgut, S., Hoppe, A., & Sonntag, K. (2016, 2016/09/02). Challenge and threat emotions as antecedents of recovery experiences: findings from a diary study with blue-collar workers. *European Journal of Work and Organizational Psychology, 25*(5), 674-689. https://doi.org/10.1080/1359432X.2015.1128414

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion, 14*(3), 573.

Monroe, A. E., & Plant, E. A. (2019). The dark side of morality: Prioritizing sanctity over care motivates denial of mind and prejudice toward sexual outgroups. *Journal of Experimental Psychology: General, 148*(2), 342.

Murray, D. R., Kerry, N., & Gervais, W. M. (2019). On Disease and Deontology: Multiple Tests of the Influence of Disease Threat on Moral Vigilance. *Social Psychological and Personality Science, 10*(1), 44-52. https://doi.org/10.1177/1948550617733518

Pizarro, D., Inbar, Y., & Helion, C. (2011, 2011/07/01). On Disgust and Moral Judgment. *Emotion Review, 3*(3), 267-268. https://doi.org/10.1177/1754073911402394

Reynolds, C. J., & Conway, P. (2018). Not just bad actions: Affective concern for bad outcomes contributes to moral condemnation of harm in moral dilemmas. *Emotion, 18*(7).

Rhee, JJ, Schein, C, Bastian, B. The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Soc Personal Psychol Compass*. 2019; 13:e12511.

Rozin, P., Haidt, J., & McCauley, C. R. (2008). *Disgust.* In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (p. 757–776). The Guilford Press.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008, 2008/08/01). Disgust as Embodied Moral Judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096-1109. https://doi.org/10.1177/0146167208317771

Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, *41*(8), 1147-1163.

Skitka, L. J., Wisneski, D. C., & Brandt, M. J. (2018). Attitude Moralization: Probably Not Intuitive or Rooted in Perceptions of Harm. *Current Directions in Psychological Science, 27*(1), 9-13. https://doi.org/10.1177/0963721417727861

Thórisdóttir, H., & Jost, J. T. (2011). Motivated closed-mindedness mediates the effect of threat on political conservatism. *Political Psychology, 32*(5), 785-811.

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology, 73*(1), 63.

Tops, M., & Boksem, M. A. S. (2011). Cortisol involvement in mechanisms of behavioral inhibition. *Psychophysiology, 48*(5), 723-732. https://doi.org/https://doi.org/10.1111/j.1469-8986.2010.01131.x

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: evolved function and structure. *Psychological review*, *120*(1), 65.

Wakslak, C. J., Jost, J. T., Tyler, T. R., & Chen, E. S. (2007). Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychological Science, 18*(3), 267-274.

Wright, J. C., & Baril, G. L. (2013). Understanding the role of dispositional and situational threat sensitivity in our moral judgments. *Journal of Moral Education, 42*(3), 383-397.

Yancey, J. R., Vaidyanathan, U., & Patrick, C. J. (2015, 2015/11/01/). Aversive startle potentiation and fear pathology: Mediating role of threat sensitivity and moderating impact of depression. *International Journal of Psychophysiology, 98*(2, Part 2), 262-269. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2014.10.014

Yoder, K. J., & Decety, J. (2020, 2020/05/18). Conflicting influences of justice motivations on moral judgments. *Cognition and Emotion, 34*(4), 670-683. https://doi.org/10.1080/02699931.2019.1669536

Zheng, Y., Xu, J., Jia, H., Tan, F., Chang, Y., Zhou, L., Shen, H., & Qu, B. (2011). Electrophysiological correlates of emotional processing in sensation seeking. *Biological psychology, 88*(1), 41-50.

Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences, 7*(3), 413-434.

Table 1.

Means, Standard Deviations and Bivariate Correlations for the relationship between Sensation-Seeking, BIS, and Social conservativism on Moral

Condemnation, as well as partial correlations for Sensation-Seeking and BIS controlling for Social Conservativism.

	M	*SD*	General Moral condemnation	Harm	Fairness	Liberty	Authority	Loyalty	Sanctity
Sensation-Seeking	13.54	6.10	-.43***	-.26**	-.31***	-.17*	-.33***	-.35***	-.39***
Partialing out social conservativism			*-.38***	*-.28**	*-.29***	*-.21*	*-.26**	*-.28**	*-.33***
Behavioral Inhibition Sensitivity	2.85	.60	.22**	.23**	.18*	.19*	.10	.14	.17*
Partialing out social conservativism			*.23**	*.24**	*.18*	*.19*	*.11	*.16†*	*.19*
Social conservativism	3.42	1.95	.25**	.00	.10	-.08	.33***	.36***	.30**

*p < .05, **p < .01, ***p < .001, †p < .10

Table 2.

Means, Standard Deviations and Bivariate Correlations for the relationship between Sensation-Seeking, BIS, Perspective Taking, Empathic Concern and Social conservativism on Moral Condemnation.

	M	SD	Perspective Taking	Empathic Concern	Social Conservativism	General Moral condemnation	Harm	Fairness	Liberty	Authority	Loyalty	Sanctity
Sensation-Seeking (0-40)	14.28	5.79	-.10	-.11	-.13	-.24**	-.25***	-.19**	-.05	-18*	-.17*	-.28***
Behavioral Inhibition Sensitivity (*BIS*)	2.82	.69	.03	.06	-.23**	.24**	.35***	.20**	.20***	.16*	.02	.18*
Perspective Taking	2.24	.51		.54***	-.01	.24**	.28**	.23**	.14*	.15*	.18**	.06
Empathic Concern	1.98	.46			.12	.28**	.26**	.22**	.19**	.25**	.27**	.12
Social Conservativism	3.50	1.85				.17*	.02	.05	-.04	.19**	.25**	.24**

*p < .05, **p < .01, ***p < .001, †p < .10.

Table 3.

Hierarchical Regression Models for Trait Sensation-Seeking predicting moral condemnation controlling for Social Conservativism, Empathic Concern and Perspective Taking.

Predictors	Outcome Variables						
	General Moral condemnation	Harm	Fairness	Liberty	Authority	Loyalty	Sanctity
Block 1							
Sensation-Seeking	-.24**	-.23***	-.17*	-.05	-.15*	-.16*	-.26***
R^2_{change}	.06**	.06**	.03*	.01	.02*	.02*	.07***
F_{change}	11.86	11.86	6.17	.45	4.43	5.08	14.47
Block 2							
Sensation-Seeking	-.19**	-.20**	-.14*	-.03	-.10	-.10	-.22**
Social Conservativism	.13*	.02	.02	-.06	.15*	.21***	.20**
Empathic-Concern	.18*	.14†	.13	.17*	.20*	.19*	.07
Perspective-Taking	.12	.18*	.15†	.05	.04	.07	.01
R^2_{change}	.09***	.08**	.06**	.04*	.08**	.11***	.05*
F_{change}	7.16	6.15	4.24	3.00	5.86	8.42	3.55
Final R^2	.15***	.13***	.09**	.04†	.10***	.13***	.11***

*p < .05, **p < .01, ***p < .001, †p < .10

Table 4.

Hierarchical Regression Models Trait BIS predicting moral condemnation controlling for Social Conservativism, Empathic Concern and Perspective Taking.

Predictors	General Moral condemnation	Harm	Fairness	Liberty	Authority	Loyalty	Sanctity
			Outcome Variables				
Block 1							
Behavioral Inhibition Sensitivity (*BIS*)	.24**	.35***	.20**	.20**	.16*	.03	.18*
R^2_{change}	.06**	.12***	.04**	.04**	.03*	.01	.03*
F_{change}	12.48	28.71	8.54	8.84	5.33	.17	6.73
Block 2							
Behavioral Inhibition Sensitivity (*BIS*)	.28***	.36***	.21**	.19**	.20**	.07	.24**
Social Conservativism	.22**	.10	.09	-.01	.21**	.24**	.29***
Empathic-Concern	.17*	.12	.11	.16†	.19*	.19*	.06
Perspective-Taking	.14†	.20**	.16*	.05	.05	.08	.03
R^2_{change}	.13***	.09***	.07**	.04†	.10***	.13***	.09
F_{change}	10.49	7.94	5.23	2.61	7.77	9.80	6.66***
Final R^2	.19***	.22***	.11***	.08**	.13***	.13***	.12***

$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$, $^†p < .10$.

Table 5.

Means, Standard Deviations and Bivariate Correlations for Variables in Study 3

	M	SD	Threat	Disgust	Harmful	Moral Condemnation
Harm Violations	3.51	.88				
Sensation-Seeking	15.15	7.08	-0.13**	-0.08	-0.11†	-0.13*
BIS	2.90	.70	0.26**	0.17**	0.24**	0.21*
Social Conservativism	3.32	1.82	.08	.04	-.06	-.04
Threat-Emotion	4.01	1.29		.74***	.65***	.71**
Disgust	4.28	1.68			.66***	.69***
Harm	4.92	1.41				.63***
Sanctity Violations	3.90	.94				
Sensation-Seeking	15.15	7.08	-0.23**	-0.25*	-0.24*	-0.24*
BIS	2.90	.70	0.21**	0.13*	0.00	0.10†
Social Conservativism	3.32	1.82	.17**	.22***	.25**	.28**
Threat-Emotion	4.91	1.26		.73***	.62***	.74***
Disgust	5.76	1.19			.48***	.73***
Harm	4.07	1.68				

*$p < .05$, **$p < .01$, ***$p < .001$, †$p < .10$.

Figure 1.

Standardized Coefficients and Indirect effects (5000 bootstrap samples) of Behavioral Inhibition Sensitivity (BIS) (Panel 1) and Sensation-Seeking (Panel 2) on Moral Condemnation of Harm (Panel A) and Sanctity (Panel B) violations via Self-Oriented Threat. Each Panel individual panel represents separate Model analyses, where social conservativism was entered as a covariate.
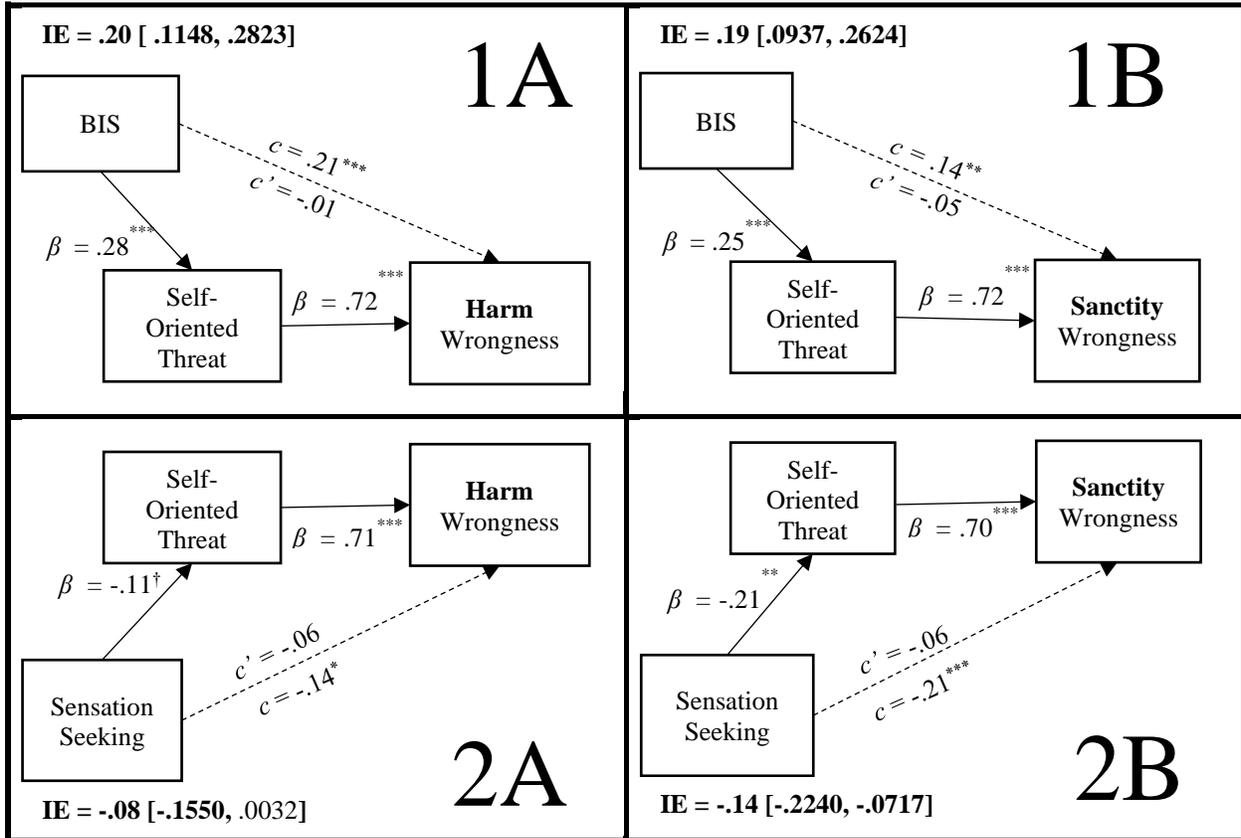
Table 6.

Total and Direct effects of Behavioral Inhibition Sensitivity (BIS) and Sensation Seeking (SS) on

Moral Condemnation of Harm and Sanctity Violations; and Indirect Effects via Threat, Harm

and Disgust, controlling for social conservativism. All mediator variables were entered

simultaneously as parallel mediators.

			Indirect effects		
	Total Effect	Direct Effect	Self-Oriented Threat [95%CI]	Perceived Harm [95%CI]	Disgust [95%CI]
BIS → Condemnation of Harm	.21***	.01	IE = 10 [.0526, .1626]	IE = .04 [.0114, .0910]	IE = .05 [.0138, .0971]
BIS →Condemnation of Sanctity	.17**	-.02	IE = .10 [.0443, .1700]	IE = .01 [-.0058, .0151]	IE = .06 [.0144, .1253]
SS →Condemnation of Harm	-.14*	-.06	IE = -.04 [-.0881, .0026]	IE = -.02 [-.0575, .0012]	IE = -.02 [-.0655, .0136]
SS →Condemnation of Sanctity	-.21***	-.03	IE = -.08 [-.1430, -.0349]	IE = -.01 [-.0368, .0033]	IE = -.08 [-.1405, -.0384]

*p < .05, **p < .01, ***p < .001, †p < .10

4

Chapter 4

An Ecological Momentary Assessment of the role of Incidental Affect on Moral Judgment in Day-to-Day Life

Chapter 4 An Ecological Momentary Assessment of the role of Incidental Affect on Moral Judgment in Day-to-Day Life

4.1.   Preface

The aim of Chapter 4 was to investigate the dynamics of within-person changes in moral judgment. Specifically, the aim of the study was to test whether people's real-world experience of affective states predicted observable changes in their tendency for moral condemnation. The manuscript presented in this chapter is in preparation for submission to *Nature Human Behaviour* and is presented in the journal submission format.

In Chapter 3 of this thesis, we found across three studies that individual differences in threat sensitivity were associated with the tendency to morally condemn third-party actions. Further, in Study 3 of Chapter 3, we found that this relationship between threat-sensitivity and tendency for moral condemnation was mediated by participants' experience of threat-appraisal emotions (i.e. anxious, worried, troubled). Collectively, these findings suggested that people who are more likely to perceive threats in their environment are also more likely to invoke morality in the evaluation of others' actions. However, an important consideration in the evaluation of the action-based moralization model is that it is fundamentally concerned with changes occurring within the individual. Indeed, by definition, moralization refers to within person changes in the degree to which moral significance is attributed to actions, attitudes or objects (Brandt et al., 2015; Rhee et al., 2019; Wisneski et al., 2020). For this reason, the development of any novel account of moralization ultimately depends on whether the proposed dynamics are observable in terms of changes occurring within the individual.

Chapter 4 thus sought to investigate the dynamics of within-person changes in moral judgment. To this end, we implemented a novel Ecological Momentary Assessment (EMA) approach – a longitudinal method which involves administering several short surveys per day over several days. This approach allowed us to investigate potentially subtle changes in participants' tendency to moralize their judgments of third-party actions with a high degree of temporal resolution. More generally, this approach also afforded us an opportunity to bring greater clarity to current understanding on the degree to which moral judgments could be *intuitive* in real-world contexts. More specifically, the approach allowed us to investigate whether participants' real-world experience of certain emotions could influence their moral judgments of third-party actions that were unrelated to the source of their emotional experience.

The specific aims of this Chapter were twofold. The primary aim of this Chapter was to clarify current uncertainties in the literature on whether certain emotions can *incidentally* influence moral judgments, by investigating whether momentary changes in participants' experience of these emotions in the real-world predicted the harshness of their moral judgments. To this end – reflecting the current literature (see: Landy & Piazza, 2019) – we focused on a range of negative emotions, which notably included disgust. The secondary aim of this Chapter was to investigate the relevance of threat-appraisal as a relevant construct in the relationship between real-world emotional experiences and moral judgment. To this end, we looked specifically at whether participants' experience of the threat-appraisal emotions measured in Study 3 of Chapter 3 predicted changes in their moral judgments.

4.2.   Manuscript

Abstract

Early influential accounts of morality suggested that so-called *incidental* disgust – that is, the experience of disgust that is completely irrelevant to the judgment target – plays an important role in shifting people's moral judgments. However, in the face of recent evidence challenging the replicability of incidental disgust findings, new perspectives have gained prominence, which both challenge the specific importance of disgust (c.f. other negative emotions), as well as the assertion that *incidental* emotions (as opposed to emotions specifically related to the judgment target) are sufficient to influence the way people make moral judgments. We investigated the relative merit of each of these perspectives using an ecological momentary assessment approach. Student participants recruited at a large Australian University ($N = 174$) took part in the seven-day study, where they were prompted six (6) times per day on their Smartphones to complete a brief survey measuring their momentary experiences of various negative emotions. At the end of each brief survey, participants were presented with a single vignette portraying a potential moral violation and were asked to rate its moral wrongness. Results provided support for the influence of incidental negative emotions on moral judgment. Specifically, we found that momentary changes in the real-world experience of negative emotions predicted moral judgments that had no specific connection to the emotional experiences.

An Ecological Momentary Assessment of the role of Incidental Affect on Moral

Judgment in Day-to-Day Life

"Since morals, therefore, have an influence on the actions and affections, it

follows, that they cannot be deriv'd from reason: and that because reason alone, as we

have already proved can never have any such influence. Morals excite passions, and

produce or prevent actions. Reason of itself is utterly impotent in this particular. The

rules of morality, therefore are not conclusions of our reason.

– Hume in *Treatise on Human Nature* Book III –

Whether it is an offensive post on social media, a news story about a corrupt politician, or simply a passerby who throws their rubbish onto the curb, we are often confronted with situations that call on us to make moral judgments about the actions of others (Hofmann et al., 2014). Such moral judgments have important implications for the way we interact with others. Judging an action as immoral can lead to perceptions that the perpetrator is untrustworthy (Pizarro & Tannenbaum, 2012), and reduce our sense of empathy toward them (Cui et al., 2016; Monroe & Plant, 2019). Despite these important implications, prevailing psychological theories suggest that moral judgments tend to be made automatically based on intuition, and can be shifted through relatively arbitrary changes in emotional experiences (Haidt, 2001). However, the empirical evidence supporting such accounts are surprisingly limited, and have, to date, almost exclusively been studied *via* analysis of questionnaire data or in highly controlled laboratory experiments designed to induce a specific emotion in isolation. Using an ecological momentary assessment approach administered *via* smartphone, we sought to investigate whether

incidental changes in participants' day-to-day emotional experiences influenced their moral judgments.

Incidental Emotion and Moral Judgment

The 18th century Scottish philosopher David Hume famously rebelled against the orthodoxy of the time that God had ordained man with the ability to reason himself to moral truth, by arguing that the very thing that defined morals were the passions they excited, and the actions they produced or prevented – both of which lay outside the domain of logic or reasoning (Hume, 1739). Thus, Hume famously proclaimed that morality was "a slave of the passions". More than two-hundred and fifty years later, Hume's observations on morality inspired a similar movement among researchers trying to make sense of the psychological processes which led to moral beliefs and decision making. Specifically, these moral psychology researchers advocated for a shift in focus from deliberative reasoning to emotions and intuition as the driving mechanism of moral judgments (Bloom, 2010; Haidt, 2007; Haidt et al., 1993). An important question arising from this shift in focus has been how much the experience of emotion can influence moral judgment - in other words, to what extent morality can be considered a "slave of the passions"?

While it is now relatively uncontroversial that emotions accompany moral judgments (Avramova & Inbar, 2013), some accounts of morality have argued that emotions play a more central role. The most influential of these has been the Social Intuitionist Model (Haidt, 2001) which argues that moral judgments are largely made automatically based on intuitions or "gut-feelings" that are partly innate and partly acquired through socialization. A core premise of such intuition-based accounts is that individuals often unconsciously use their experience of certain emotions as guides when making moral judgments. Perhaps the most well-researched emotion in

the context of intuition-based moral judgment is disgust (e.g. Rozin et al., 2008; Tybur, Lieberman, Kurzban, & DeScioli, 2013). Early studies focused on the influence of *incidental disgust* on moral judgments – more specifically, whether disgust elicited by one event could shape later judgments that were unrelated to the emotion-eliciting event (Horberg et al., 2009). For example, notable early studies used stimuli such as fart spray (Schnall et al., 2008), bitter drinks (Eskine et al., 2011), and hypnosis (Wheatley & Haidt, 2005) to induce disgust in participants, and subsequently found that this increased the severity of participants' unrelated moral judgment about another person's actions. Such findings not only established the relevance of disgust to moral judgments, but also supported the notion that moral judgments could occur intuitively based on individuals' experience of certain emotions at the time of the judgment.

However, there is now mounting evidence challenging the replicability of *incidental* disgust findings (see: Landy & Goodwin, 2015; Ghelfi et al., 2020). For example, a meta-analysis conducted by Landy and Goodwin (2015) found only a small effect of incidental disgust inductions on increased moral condemnation, which became non-significant once publication bias was accounted for. Similarly, in a recent many labs replication project, Ghelfi et al. (2020) found little evidence to support previous findings by Eskine et al. (2011) that disgust-inducing drinks led to more severe moral judgments. Given that past studies linking *incidental* emotions to moral judgment focused almost exclusively on inductions of disgust, the questioning of such evidence not only challenges the specific role of *incidental* disgust in moral judgment, but also the premise that moral judgments are shaped by incidental emotions in general.

The Current Study

In the current study, we sought to investigate whether *incidental* emotions can indeed influence moral judgment in an ecologically valid setting. One potential reason for the

inconsistency in findings regarding emotions and moral condemnation may be the artificiality of the environments in which such emotions are induced. As noted above, most incidental affect studies have either been conducted via online surveys, or in laboratory settings, using highly unusual and evocative stimuli (e.g. fart-spray, images of dirty rooms, bitter drinks: see Landy & Goodwin, 2015). While such studies provide insight into whether emotions *can* influence moral judgment, they leave important gaps in understanding about whether such emotions are actually relevant in peoples' day-to-day moral judgments. Furthermore, past studies on the relationship between incidental emotions and moral judgment have focused specifically on the emotion of disgust, despite mounting evidence that a number of other negative emotions may be relevant to moral judgment (Cheng et al., 2013; Landy & Piazza, 2019). Finally, while studies such as those conducted by Landy & Piazza (2019) suggest that disgust may not play a special role in moral judgment, it remains unclear whether the relationship between negative emotions and moral judgement can be understand through a different functional lens.

To address these gaps in understanding, we conducted the first study to date employing an experience sampling method to investigate whether momentary changes in the experience of certain negative emotions predict subsequent moral judgments of actions that are completely unrelated to the source of the emotion. Participants were asked six times per day to indicate their current experience of a range of negative emotions, and were then asked to rate the moral wrongness of a randomly selected vignette. Given that the vignettes were presented to participants after they had indicated their current emotional experience, it was possible to infer that any relationship between changes in emotional experience and subsequent moral judgment could only occur *incidentally*. Finally, in order to explore potential functional themes in any observable relationship between incidental affect and moral judgment, we also measured threat-

appraisal emotions (Michel et al., 2016) alongside the negative emotions measured by Landy &

Piazza (2019). All analyses reported in the main results were pre-registered and are accessible at:

https://osf.io/wm2p8.

Methods

Participants

A total of one-hundred and eighty-six participants took part in both the pre-survey and

SEMA protocol. Recruitment took place between October 31st, 2020 and December 30th, 2020.

Participant recruitment was conducted in two waves, in accordance with our pre-registered

recruitment protocol. First, student participants ($n = 81$) at a large Australian University took

part in the study for course credits. The remaining participants ($n = 105$) were recruited via a

paid research participation program based at the same University. Participants recruited in the

second wave were reimbursed with a AU$50 gift voucher if they completed at least 80% (33 out

of 42) of the momentary assessment surveys.[1] Based on our pre-registered exclusion criteria,

participants were excluded from all analyses if they failed to complete 80% of the momentary

assessment surveys during the 7-day period. After all exclusions, a total of one-hundred and

seventy-one ($N = 171$) participants were retained for analyses.

Materials

Momentary Negative Emotions. Participants' momentary experience of emotions was

measured using an adapted version of the 10-item negative emotion terms employed by ***Landy &***

Piazza (2019). The negative emotions were presented in randomized order, and participants were

---

[1] Participants were informed of the 80% participation hurdle requirement for reimbursement prior to their
participation in the study. Participants who failed to meet the 80% requirement, but completed at least 70% of the
surveys were reimbursed with a AU$30 gift voucher for their participation.

asked to indicate how much they currently felt each emotion on a seven-point scale (1 = Not at all, 7 = Very much). The two items measuring disgust ("grossed out", "disgusted"), were averaged to form a disgust-composite where higher scores indicated higher experience of disgust, R_{KF} = 1.00; R_C = .74. The remaining eight items ("anxious", "angry", "irritated", "distressed", "negative", "alert", "calm", "afraid") were used as a measure of general negative emotion. Two of these items were negatively scored ("alert", "calm"), and all items were averaged such that higher scores indicated higher experience of negative emotions, R_{KF} = .99; R_C = .49. All measures of disgust, and negative emotions were participant mean-centered to create within-person effects (i.e., participant's mean was subtracted from each of their scores).

Momentary Threat-Appraisals. Participants' self-oriented threat-appraisals in response to the vignettes were assessed using the threat-appraisal emotions used by (Michel et al., 2016). Specifically, participants were asked to indicate the extent to which they felt "worried", "anxious" and "troubled". The items were averaged to form a threat-appraisal emotion composite such that higher scores indicated higher experience of t threat-appraisal emotion R_{KF} = 1.00; R_C = .81. All measures of threat-appraisal emotions were participant mean-centered to create within-person effects (i.e., participant's mean was subtracted from each of their scores).

Moral Vignettes. Participants' momentary moral judgments were measured using a subset of the 117 vignettes compiled by Kruepke et al. (2018) and first developed by Knutson et al. (2010). The content of these vignettes were originally derived by asking participants to write a short recollection of a time they engaged in a morally relevant episode, which were subsequently condensed by the researchers into two or three sentences (Knutson et al., 2010). Of these vignettes, we selected only those which were rated as morally inappropriate on average by a norming sample of 661 participants reported by Kruepke et al. (2018). Vignettes were also

removed if they contained references to US-specific contexts (e.g. "*A little while ago I had to go to the DMV to get my drivers license renewed.*[...]"). The final selection consisted of seventy-seven vignettes (e.g. "*When I was a teenager, I took my father's car out with my friends. My parents were out of town for a couple of days. I took the car without their permission to go for a joyride.*").

Moral Judgment. Participants' moral judgment at each observation was measured using one item, which asked how morally wrong participants' considered the person's action in the vignette they had just read. Participants responded on a 7-point scale (1 = *Not at all*, 7 = *Very much*).

Procedure

All data collection took place remotely. Participants were first sent an initial questionnaire containing a series of individual differences and demographic measures. Once they completed the initial questionnaire, participants were contacted *via* telephone or video conferencing call by one of researchers who provided them with detailed instructions on how to proceed with the momentary assessment portion of the study. All momentary assessment measures were administered using the SEMA3 platform (Koval et al., 2019). Participants were asked to download the SEMA3 application on their smartphones, and to ensure that notification permissions were enabled for the application. During the seven-day period of the study, participants received a notification six-times per day prompting them to fill out a brief questionnaire for the study. Each notification was sent at 75 minute intervals from 10am to 10pm, and participants were given a 60 minute window to complete each survey.

Within each questionnaire, participants first responded to all items measuring their current emotional experience. Participants were then presented with a single randomly selected

vignette.[2] Prior to reading the vignette, participants were informed that they were about to read a real-life account written by an actual person. After reading the vignette, participants were asked to rate the moral wrongness of the person's action portrayed therein. After the 7-day period, participants were contacted by the researchers for debriefing.

Results

Statistical Analysis

We conducted our analyses using the *lme4* package (Bates et al., 2015). We ran two-level models in which observations ($N = 42$) were nested within vignette ($N = 77$), and within persons ($N = 171$). To aid interpretability and reduce convergence issues, all variables were standardized and scaled. For the purposes of the statistical models, both disgust and negative emotion composite scales were allowed to vary randomly across vignettes and participants (Nezlek, 2011).

To test each of our hypotheses, we ran three multilevel models that examined the relationship between moral condemnation and the predictors disgust, negative emotions, and threat-appraisal emotions. In each of these three models we focused on the within-person and lagged effects of our predictors. Specifically, these predictor variables captured one's level of disgust, negative emotions, and threat-appraisal emotions at time i-1 compared to their average level. Following suggestions from Bolger and Laurenceau (2013), each model included a number of covariates. First, each model controlled for autoregression by including lagged moral condemnation scores at time i-1. This also allowed us to model *change* in moral condemnation (from i-1 to i) as a function of the predictor variable. Second, to control for variance in moral

---

[2] Randomization was without replacement. This meant that no participants saw the same vignette twice throughout the period of the study. Given that there were 42 total surveys (6 surveys/day over 7 days), no participant saw the full set of 77 vignettes during the study period.

wrongness across vignettes, we included norming-ratings of the moral appropriateness of all vignettes measured in a previous sample of 661 participants (Kruepke et al. 2018). Finally, to account for any effects associated with the passage of time, we included an *observation* variable that indexed the 42 observations of the measurement period. Results for all models are presented in Table 1[3].

---

[3] Visual inspection of diagnostic plots produced using the plot_model function from the sjPlot package in R were used to assess normality, linearity, and homoscedasticity (Lüdecke, 2020). All models satisfied these assumptions.

Table 1

Multilevel models that examining the relationship between moral condemnation and the predictors disgust (Model 1), negative emotions (Model 2), and threat-appraisal emotions (Model 3).

	Parameter	β	95% CI	*p*-value	ICC	*N* (observations)
Model 1	*Incidental Disgust*				.25	169 (5413)
	Intercept	.14	-.02, .30	.087		
	Disgust *i*-1	.01	-.10, .13	.214		
	Normed Moral Appropriateness	-.50	-.55, -.44	< .001		
	Moral Condemnation *i*-1	-.01	-.03, .01	.423		
Model 2	*Incidental Negative Affect*				.25	169 (5413)
	Intercept	.14	-.02, .30	.092		
	Negative Affect *i*-1	.02	.00, .04	.023		

		β	CI	p	ICC	N
	Normed Moral Appropriateness	-.50	-.55, -.44	< .001		
	Moral Condemnation i-1	-.01	-.03, .01	.388		
Model 3	*Threat-Appraisal Emotion*				.25	169 (5413)
	Intercept	.14	-.02, .30	.093		
	Threat-Appraisal Emotion i-1	.02	.00, .04	.021		
	Normed Moral Appropriateness	-.50	-.55, -.44	< .001		
	Moral Condemnation i-1	-.01	-.03, .01	.384		

Note. β = Standardized beta; CI = Standardized confidence interval; ICC = Intra-class correlation; *i-1* = Score from the previous time point; Normed Moral Appropriateness = Average moral ratings normed on a sample of 661 participants (see: Kruepke et al., 2018).

The "observation" variable is omitted from each model for ease of interpretation

Incidental Disgust Hypothesis

Model 1 was used to test the incidental disgust hypothesis, where lagged disgust was used to predict change in moral condemnation. As shown in Table 1, we found that lagged disgust did not significantly predict changes in moral condemnation.

Incidental Negative Emotion Hypothesis

Model 2 was used to test the incidental negative affect hypothesis where lagged negative affect was used to predict change in moral condemnation. As shown in Table 1, we found that lagged negative affect significantly predicted changes in moral condemnation. This effect suggests that when participants experienced greater than average negative emotions at a given time point, they tended to increase their moral condemnation at the next time point.

Threat-Appraisal Emotion Hypothesis

Model 3 was used to test the threat-appraisal emotion hypothesis. As shown in Table 1, we found that lagged threat-appraisal emotion significantly predicted moral condemnation.

Given the conceptual overlap between negative affect and threat-appraisal emotions explored in this study, we conducted a follow-up exploratory factor analysis to investigate whether the predictive relationship between these measures and moral condemnation could reflect a common underlying construct. This analysis is presented in the **supplementary analyses**.

<div align="center">

Discussion

</div>

Our findings show that momentary changes in participants' day-to-day experience of negative emotions incidentally increase the harshness of their moral judgments. Such findings echo the views of influential philosophers and psychologists who have long theorised that moral judgments can often be driven by emotion and intuition over reason and logic. They also provide

the first empirical evidence to date that this incidental effect of emotion on moral judgment is observable in people's natural environments.

Interestingly, while our findings support past theories suggesting that moral judgments can be intuitively impacted by people's affective states (Haidt, 2001; 2007), we did not find this effect for the traditionally "moral emotion" of disgust. We did however, find that negative emotions associated with threat-appraisal ("worried", "troubled" and "anxious") did predict moral judgments, indicating the potential value of further exploring the relationship between threat-appraisal emotions and moral judgement.

While the design of our study certainly does not rule out the relevance of traditional moral emotions to people's moral judgments, they do suggest that certain dynamics of moral judgment occurring in people's natural environments may be less readily observable in controlled experimental settings. These revelations highlight the potential value of ongoing research applying ecologically valid methods to evaluate the psychology of moral judgments.

References

Bloom, P. (2010, 03/24/online). How do morals change? *Nature, 464*, 490.
https://doi.org/10.1038/464490a

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A Constructionist Review of Morality and
Emotions:No Evidence for Specific Links Between Moral Content and Discrete
Emotions. *Personality and Social Psychology Review, 19*(4), 371-394.
https://doi.org/10.1177/1088868314566683

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral
judgment. *Psychological Review, 108*(4), 814-834. https://doi.org/10.1037/0033-
295X.108.4.814

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science, 316*(5827), 998-1002.
https://doi.org/10.1126/science.1137651

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat
your dog? *Journal of Personality and Social Psychology, 65*(4), 613.

Horberg, E. J., Oveis, C., & Keltner, D. (2011). Emotions as Moral Amplifiers: An Appraisal
Tendency Approach to the Influences of Distinct Emotions upon Moral Judgment.
Emotion Review, 3(3), 237-244. https://doi.org/10.1177/1754073911402384

Landy, J. F., & Goodwin, G. P. (2015, Jul). Does Incidental Disgust Amplify Moral Judgment?
A Meta-Analytic Review of Experimental Evidence. *Perspect Psychol Sci, 10*(4), 518-
536. https://doi.org/10.1177/1745691615583128

Landy, J. F., & Piazza, J. (2019). Reevaluating Moral Disgust: Sensitivity to Many Affective
States Predicts Extremity in Many Evaluative Judgments. *Social Psychological and
Personality Science, 10*(2), 211-219. https://doi.org/10.1177/1948550617736110

Nelissen, R. M., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party
punishment: Anger, guilt and the functions of altruistic sanctions. *Judgment and Decision
Making, 4*(7), 543.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: a mapping
between three moral emotions (contempt, anger, disgust) and three moral codes
(community, autonomy, divinity). *Journal of Personality and Social Psychology, 76*(4),
574.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008, 2008/08/01). Disgust as Embodied
Moral Judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096-1109.
https://doi.org/10.1177/0146167208317771

Wheatley, T., & Haidt, J. (2005). Research Report: Hypnotic Disgust Makes Moral Judgments More Severe. *Psychological Science, 16*(10), 780-784. https://doi.org/10.1111/j.1467-9280.2005.01614.x

Supplementary Analyses

The pre-registered analyses reported in the main results revealed the following:

1. Disgust did not predict momentary changes in moral judgement;

2. A composite scale of negative affect used by Landy & Piazza (2017) predicted momentary changes in moral judgment; and

3. A composite scale of threat-appraisal emotions predicted momentary changes in moral judgment.

The finding that both the negative affect and threat-appraisal composite scales predicted momentary changes in moral judgment provides support for the relevance of incidental emotions on moral judgement. However, given the conceptual overlap between negative affect and threat-appraisal emotions, the nature of the underlying mechanism that links these constructs to moral judgment remains unclear. More specifically, it remains unclear from the primary analyses whether the observed relationship between both incidental negative and threat-appraisal emotions, and moral judgment could be explained by the negative valence alone, or by an alternative latent construct that sits within the experience of negative affect.

To disambiguate these possible alternative explanations for the main results, we conducted a preliminary exploratory factor analysis (EFA) using principal axis factoring (PAF) across all emotion items used to measure negative affect and threat-appraisal emotions. We then created new composite affect scales based on the factor structure identified in the EFA. These new affect scales were then entered as lagged predictors in the mixed-effects models predicting moral condemnation used in the main analyses. If negative valence was the latent construct accounting for the relationship predictive relationship between moral judgment and incidental experience of both negative affect and threat-appraisal emotions, we would expect any factors identified in the

EFA to be predictive of moral condemnation. Alternatively, any differential pattern of relationship between factors identified in the EFA and moral condemnation would suggest there was an alternative latent construct sitting within the experience of negative affect that was relevant to moral judgment.

Results

Factor Analysis

Factor Structure. Exploratory factor analysis (EFA) using principal axis factoring was conducted using the Oblimin method of oblique rotation. Two factors were extracted by applying Kaiser's criterion of Eigenvalues of ≥1 (see Table 2). Collectively these two factors accounted for 48% of variance. One item ("Alert") was dropped due to low loadings on both factors (<.30).

Factor 1 (Anxious-threat) Factor 1 included six (6) items in total, including all three items used to measure threat-appraisal emotions as well as anxiety-related items from the negative affect scale including "Distressed", "Calm", and "Afraid".

For the purposes of further modelling, one item was reverse scored ("Calm"), then all items were averaged to form a composite where higher scores indicated higher experience of "anxious-threat" emotions, $M = 2.65$, $SD = 1.10$, $\alpha = .81$.

Factor 2 (Anger). Factor 2 included two (2) items, including "Angry", and "Irritated". For the purposes of further modelling, the two items were averaged to form a composite where higher scores indicated higher experience of "angry" emotions $M = 1.97$, $SD = 1.01$, $r = .84$.

Table 2.

Pattern Factor of Loadings for Exploratory Factor Analysis on Negative Affect and Threat-Appraisal Emotions (N = 7326).

Item No.	Items	Factor 1	Factor 2
1	Angry		**.82**
2	Irritated		**.71**
3	Distressed	**.53**	.33
4	Alert*		.14
5	Calm	**-.47**	-.11
6	Afraid	**.48**	.33
7	Worried	**.93**	
8	Anxious	**.86**	
9	Troubled	**.74**	.10
	Eigenvalues	2.90	1.44
	% of Variance	32.3	48.2

Note. Factor loadings are in boldface. *The Alert item was dropped due to insufficient loading on both factors.
Extraction Method: Oblimin rotation

Multi-level modelling

To investigate the relevance of the two factors extracted in the EFA to moral condemnation, we replicated the multi-level analyses conducted in the main analyses, using each of the new two factors as predictor variables. Specifically, we conducted three separate models focusing on the lagged effects of participants' level of "anxious-threat", and "anger" on moral condemnation. As with the main analyses, each model controlled for lagged moral condemnation scores at time i-1, norming-ratings of the moral appropriateness of all vignettes measured, and an observation variable that indexed the 42 observations of the measurement period.

Model 1 was used to investigate the predictive relationship between the "anxious-threat" factor and moral condemnation. We found that lagged "anxious-threat" factor emotions significantly predicted moral condemnation, $\beta = .02$, $p = .033$, 95% CI [.00, .04].

Model 2 was used to investigate the predictive relationship between the "anger" factor and moral condemnation. We found that lagged "anger" factor emotions did not predict moral condemnation, $\beta = .01$, $p = .033$, 95% CI [-.02, .03].

Finally, in Model 3, we investigated whether the "anxious-threat" factor would continue to predict moral condemnation even after controlling for the "anger" factor. When entered into the same model, we found that lagged "anxious-threat" factor emotions significantly predicted moral condemnation ($\beta = .02$, $p = .038$, 95% CI [.00, .05]), while the "anger" factor remained a non-significant predictor, $\beta = .005$, $p = .731$, 95% CI [-.03, .02].

Discussion

In this supplementary analysis we aimed to disambiguate findings in the primary analyses by investigating whether the observed predictive relationship between both incidental negative and threat-appraisal emotions, and moral judgment could be explained by: (1) negative valence alone; or (2) by an alternative latent construct that sits within the experience of negative affect. Our findings provide preliminary support for the second account. Specifically, we find that negative affect and threat-appraisal items conform to a two-factor structure that could be described as "anxious-threat" and "anger" constructs. Further, we find that only the "anxious-threat" construct is a significant predictor of moral condemnation.

The findings of this supplementary analysis lend further support to the relevance of threat-appraisal in people's moral judgments. However, it is important to note that these analyses were highly exploratory, and were not part of the original pre-registered analysis plan. These

findings should therefore be interpreted with caution, and future research should seek to validate these findings through specific *a priori* testing. Future research should specifically prioritise validation of the "anxious-threat" factor identified in the EFA via confirmatory factor analysis.

4.3. Summary of Key Findings in Chapter 3 and Chapter 4

The overarching aim across Chapters 3 and 4 of this thesis was to empirically test one of the core premises of the action-based moralization account – specifically, that individuals would be motivated to moralize in response to threat-appraisal.

4.3.1. *Consistent relationship between Threat-Appraisal and Moral Judgment across multiple domains*

In Chapter 3, we found that individuals who are higher in trait threat-sensitivity – and therefore more likely to appraise situations as threatening – were also more likely to perceive third-party actions as morally wrong across a range of moral domains. Further we found in Study 3 of Chapter 3 that this relationship between trait threat-sensitivity and tendency for moral condemnation was mediated by participants' experience of threat-appraisal emotions. In Chapter 4 we found that this between-subjects relationship between individual differences in threat-appraisal and moral condemnation could also be observed in terms of within-person changes in the experience of threat-appraisal emotions. Specifically, we found that variance in participants' momentary experiences of threat-appraisal emotions was a significant predictor of variance in their tendency to morally condemn third-party actions. Taken together, these findings provide initial empirical support for the premise of the action-based moralization account that individuals will be particularly motivated to moralize under conditions of threat-appraisal.

4.3.2. *Threat/Moralization Relationship not Accounted for by Disgust*

Disgust has traditionally taken a prominent position in theorising around moral cognition (see: Rozin, et al., 2008; Giner-Sorolla & Chapman, 2017; Tybur, et al., 2013), with accounts suggesting that the emotion's utility as a clear affective signal has led to it being evolutionarily co-opted as a distinctly moral emotion (see: Tybur et al., 2013). Given this prominence of disgust

in previous theorizing on moralization, it was important to account for the possibility that any observed relationship between threat-appraisal and changes in moral judgment could be attributed to shared variance between these constructs and participants' propensity to feel disgust.

In Study 3 of Chapter 3, we first addressed this possibility by measuring participants' experience of disgust upon reading the vignettes that they judged in terms of moral wrongness. We found in parallel mediation analyses that the indirect effect of trait threat-sensitivity on the tendency for moral condemnation *via* the experience of threat-appraisal emotions remained significant even when disgust was included as a parallel mediator to the model. In Chapter 4, we further found that momentary changes in participants' experience of disgust was not a significant predictor of changes in their moral judgments. While neither the findings in Chapter 3 or Chapter 4 can be taken as excluding the relevance of disgust to moral judgment, they do provide initial evidence that participants' the observed relationships between threat-appraisal and moral judgment cannot be fully accounted for by the construct of disgust.

4.3.3. *Threat-Appraisal Incidentally Affects Moral Judgment*

In Chapter 4, we found support for the notion that the experience of threat-appraisal emotions could *incidentally* influence moral judgment. Specifically, we found that momentary changes in the real-world experience of threat-appraisal emotions predicted moral judgments of third-party actions that had no specific connection to the emotional experiences. Such findings appear consistent with influential accounts of morality that have suggested moral judgments are often the product of automatic and intuitive processes (e.g. Haidt, 2001). However, we surprisingly found that emotions that have traditionally been conceptualised as morally relevant (i.e., disgust and anger) did not predict momentary changes in moral judgment. The potential

implications of these findings in terms of theorising on the mechanisms by which emotions intuitively influence moral judgments will be considered in the General Discussion of this thesis.

Chapter 5
Threat-Appraisals are Positively Related to
Moralization of COVID-19 Behaviours and
Desire for Tight Cultural Norms

Chapter 5 COVID-19 Threat-Appraisals are Positively Related to Moralization of COVID-19 Behaviours and Desire for Tight Cultural Norms

5.1.   Preface

The aim of the current study was to investigate whether the threat-appraisal/moralization relationship could also be observed in a relatively novel context – specifically that of the COVID-19 pandemic. The manuscript presented in this chapter was submitted for publication in *Proceedings of the Royal Society: B* and is presented in author-submitted format.

Across Chapter 3 and 4, we found increasing support for the general relationship between threat-appraisal and moralization of judgments. In Chapter 5, we sought to further evaluate the robustness of this relationship by investigating whether threat-appraisal/moralization dynamics could also be observed in participants' judgments of relatively novel actions. To this end, we used the context of the SARS-CoV2 ("COVID-19") pandemic – a once in a generation event that was accompanied by the widespread prescription of relatively novel behaviours required to reduce the spread of the COVID-19 virus. In this study, we recruited two representative samples (one from the U.S. and U.K) to investigate whether individual differences in threat-appraisal regarding the COVID-19 pandemic were related to the tendency to morally condemn others' failures to engage in these novel COVID-19 preventing behaviours.

5.2.   Manuscript

COVID-19 Threat-Appraisals are Positively Related to Moralization of COVID-19 Behaviours and Desire for Tight Cultural Norms

Joshua J. Rhee[1], Michael R. Donner[1], Christoph Klebl[1], Melanie J. McGrath[1], Ji Young Song[1], Nicholas Poh-Jie Tan[1], Morgan Weaving[1], Chiara Van Reyn[2] & Brock Bastian[1]

[1] School of Psychological Sciences, University of Melbourne, Parkville VIC, Australia, 3010

[2] Katholieke Universiteit Leuven

Author Note

Correspondence: Joshua J. Rhee, Melbourne School of Psychological Sciences, University of Melbourne, Victoria, 3010, Australia; email: rheej@student.unimelb.edu.au

Declarations

Abstract

Tight-Loose culture theory proposes that cultures can develop strong norms and a low tolerance of deviant behaviour (cultural "tightness") as a way of adapting to environments requiring high levels of coordination to overcome ecological threats. However, little is known about how the link between ecological threats and development of tight cultures manifests at the individual-level. We investigated this dynamic in the context of COVID-19, and also investigated the potential relevance of moralization as an individual-level process accompanying the desire for cultural tightness. In representative samples drawn from the U.K. ($N = 430$) and the U.S. ($N = 830$), we investigated the degree to which self-reported threat-appraisals regarding COVID-19 were associated with perceived need for tighter social norms and moral condemnation of failures to adhere to COVID-19 mitigating behaviours (e.g., maintaining social distance). As predicted, we found significant positive correlations between COVID-19 threat-appraisal, desire for cultural tightness and moral condemnation of failing to engage in COVID-19 mitigating behaviours in both the UK and US. Further, we found the relationship between COVID-19 threat-appraisal and moral condemnation was partially mediated by a desire for cultural tightness. This study provides insight into how individual-level appraisals of COVID-19 threat may initiate changes in cultural tightness.

140

COVID-19 Threat-Appraisals are Positively Related to Moralization of COVID-19

Behaviours and Desire for Tight Cultural Norms

The 2019 novel coronavirus (COVID-19) pandemic is a global and ubiquitous threat to individuals' health and wellbeing. As of December 2021, the World Health Organisation estimates that the number of confirmed cases of COVID-19 has surpassed 290 million, resulting in more than 5.4 million deaths across almost 200 countries (*WHO Coronavirus (COVID-19) Dashboard*, 2021). A defining feature of COVID-19 is its ease of transmission. Current estimates of the pooled global basic reproduction number of COVID-19 suggest that, on average, each person contracting the virus will go on to infect approximately four others (Ganyani et al., 2020). The potential for a small number of cases to quickly surge into widespread community outbreaks makes the containment of COVID-19 a collective action problem (Ostrom, 2000) where countries require near universal cooperation by members of a given population through adoption of COVID-19 mitigating behaviors and adherence to public health guidelines (Gelfand et al., 2021; Harring et al., 2021). Reflecting this need for widespread co-ordination in effective COVID-19 mitigation, country-level comparisons strongly suggest that collectivist countries have experienced fewer COVID-19 cases and fatalities per capita than individualistic countries (Maaravi et al., 2021).

According to Tight-Loose culture theory (Gelfand et al., 2011), this tendency for collectivist cultures to be more effective in their pandemic response can be accounted for by their proclivity for cultural tightness - that is, having clear social rules and low tolerance for deviant behaviour. Tight-Loose culture theory posits that tight cultures develop in environments requiring high levels of coordination to ensure survival, such as those with high resource constraints, or high levels of threat to group members (Gelfand et al., 2011). Supporting the

notion that cultural tightness specifically facilitates group-level coordination against collective threats, Gelfand et al. (2020) found that across 59 countries, cultural tightness was a negative predictor of both COVID-19 cases and fatalities, even after controlling for country-level differences in general collectivism (i.e., ingroup favouritism and outgroup avoidance). Such findings strongly support the relevance of pre-pandemic levels of cultural tightness (Jackson et al., 2020) on COVID-19 responding.

However, it is important to note that culture is dynamic, and constantly changes as a function of interactions between individuals and their social and natural environments (Kemmelmeier & Kühnen, 2012; Markus & Kitayama, 2010). To this end, the current literature provides little empirical insight into the effect that COVID-19 may have in driving cultural change (Kashima et al., 2021). Further, the overwhelming majority of empirical work applying Tight-Loose culture theory has focused on country-level variance, meaning that little is known about how COVID-19 driven cultural change might manifest in terms of individual-level processes. Such gaps in understanding are meaningful given the mounting evidence that cultural tightness increases the effectiveness of COVID-19 mitigation, and the fact that changes in cultural tightness necessarily start at the level of individual cognition (Mrazek et al., 2013). We therefore seek to address this gap in understanding by integrating Tight-Loose culture theory with notions of threat-appraisal at the individual level. Specifically, we propose that individuals who perceive COVID-19 as an unmanageable threat may desire greater cultural tightness as a means of securing social coordination against threats. Further, we identify potential overlaps between theories of cultural tightness and recent theorising on the functions of moral judgment (DeScioli & Kurzban, 2013; Gelfand et al., 2011; Mrazek et al., 2013; Petersen, 2013), and

investigate the possibility that individuals' moralization of COVID-19 mitigating behaviours

("COVID-19 behaviours") may be partially driven by a desire for cultural tightening.

Individual Threat-Appraisal and Cultural Tightness

The core premise of Tight-Loose culture theory is that cultural tightness emerges in

response to ecological and societal threats, which increase the need for strong norms and

punishment of deviant behaviour in order to enhance social coordination (Gelfand et al., 2011;

(Roos et al., 2015). Consistent with the centrality of group-level threats in the emergence of

cultural tightness, measures of Tight-Loose at multiple levels of analysis (e.g. country, state,

regions) are consistently predicted by ecological (e.g. prevalence of disease) and human-made

threats (e.g. territorial conflict) that can only be effectively mitigated with high levels of

coordination (Gelfand et al., 2017). Thus, the starting point for understanding how cultural

tightening in response to COVID-19 may manifest at the individual level, is to consider how

such threats are evaluated by individuals.

At any given time, individuals are liable to face dangers that they need to overcome.

However, a long tradition of research suggests that whether a danger is appraised as threatening

will depend on whether the individual perceives themselves as having sufficient resources to

overcome it (Blascovich & Tomaka, 1996; Folkman et al., 1986). According to this threat-

appraisal framework, when an individual encounters a potential danger that they have sufficient

resources to overcome, they enter into an approach-oriented "challenge" state that maximises

their ability to respond effectively (Blascovich & Tomaka, 1996). On the other hand, when the

risk exceeds an individual's available resources, they perceive the situation as threatening, and

feel unable to overcome it unaided.

When faced with a situation that an individual appraises as threatening, one strategy for increasing the availability of resources to overcome a danger is to secure the cooperation of others (Blascovich, 2008; Meijen et al., 2020). This may be particularly important in situations such as the COVID-19 pandemic, where the high transmissibility of the COVID-19 virus means that individuals have relatively little control over their potential exposure to infection (Harring et al., 2021). Applying Tight-Loose culture theory to this context, it is possible that individuals may be motivated to increase cultural tightness in their society when they appraise COVID-19 as a danger to themselves that they have insufficient resources to overcome. From this perspective, the desire for cultural tightness might stem from a perception that cultural tightness would increase social cooperation against the danger of COVID-19. However, this leaves open the question of how individuals seek to realise their desire for greater cultural tightness. In the next section we propose that moralization may be one process by which this occurs.

Moralization of COVID-19 Behaviours

A growing body of research suggests that moralization—attaching moral significance to behaviors (Rhee et al., 2019; Rozin, 1999)—may play an important role in the way that individuals engage with COVID-19 behaviours (for review see: Prosser, et al., 2020). Moralization of COVID-19 behaviours has been found to predict real-world adherence to government recommended behaviours during the pandemic (Francis & McNabb, 2020), with some evidence suggesting that it may be a stronger predictor of social-distancing behaviours than empathy for others, and fear of punishment for failing to adhere to state-ordered mandates. Despite these potentially important implications of moralization in the context of COVID-19, research is yet to identify the psychological factors that may drive individuals to moralize COVID-19 behaviours.

While traditional conceptions of morality suggested that the range of actions and attitudes considered to be moral is relatively static and ubiquitous (see: Lovett & Jordan, 2010), recent research suggests that substantial variance can exist between individuals on the degree to which they moralize a given issue, object or action (Ryan, 2017; Skitka, 2010; Skitka et al., 2005). Such insights regarding the dynamic nature of moralized beliefs have led to increasing interest in *why* a given individual may choose to attribute moral significance to new beliefs or increase the moral urgency of existing beliefs. Notably, while earlier accounts of the functions of morality focused predominantly on benefits at the group-level such as altruism (Gintis et al., 2008) and discouraging selfish agents (Whitehouse et al, 2019), more recent accounts have focused on the functions that moralization may serve for the individual (see: DeScioli & Kurzban, 2013; Petersen, 2013).

Individual-focused accounts suggest that morality provides a psychological architecture for signaling to group members how social resources are to be allocated (Petersen, 2013), and who should be targets of social punishment (DeScioli & Kurzban, 2013). Importantly, these accounts posit that the content of norms that are moralized are dynamic, and are constantly renegotiated by the members of a given social group (Bor et al., 2020; DeScioli & Kurzban, 2013; Petersen, 2013). From this perspective, individual members in a given society may moralize certain norms of behaviour as a way of advocating for those norms to become the new signals for social resource allocation and social punishment (Petersen, 2013). This may be particularly adaptive for individuals who perceive themselves as having fewer personal allies or other means of protecting their personal interests (Petersen, 2013).

When viewed in light of COVID-19 and Tight-Loose culture theory, this individual focused account of moralization may provide one avenue by which individuals seek to realise

their desire for cultural tightness. Specifically, individuals who desire cultural tightness in response to COVID-19 threat-appraisals may moralize COVID-19 behavioural norms as a means of advocating that these norms be subject to stronger social regulation (i.e., a "tightening" of norms around COVID-19). This link between moralization and cultural tightness is supported by past work finding that individuals living in tight cultures are less likely to endorse the moral justifiability of norm violations (Gelfand et al., 2011; Mrazek et al., 2013). However, such work has been carried out exclusively at the group-level. As such, it has often been suggested that stricter moral condemnations for moral violations may be a consequence or manifestation of cultural tightness (Gelfand et al., 2011; Graham et al., 2016). In the present study, we sought to investigate the possibility that moralization may play a role in the process of cultural tightening, rather than simply being a consequence thereof.

The Current Study

The COVID-19 pandemic presents a unique opportunity to investigate the theorized relationships between threat-appraisal, desire for cultural tightness, and moralization in the context of a real-world threat. Such a study can be particularly powerful when combined with study samples that are representative of the general population (Hogarth, 2005). Thus, using representative samples drawn from the U.S. and U.K., the current study aimed to investigate how threat appraisals occuring at the individual level may relate to moralization of COVID-19 mitigating behaviours, and desire for cultural tightness. We predicted that individuals who were more sensitive to the health threat posed by COVID-19, and perceived themselves as having insufficient resources to deal with the threat ("high Threat-Appraisal"), would report a greater desire for cultural tightness. We also predicted that high Threat-Appraisal of COVID-19 would be positively related to moralization of COVID-19 behaviours. Finally, we predicted that the

relationship between threat-appraisal and moralization of COVID-19 behaviours would be mediated by desire for cultural tightness. All three hypotheses were pre-registered.[4]

Study 1

We first sought to test our hypotheses using a representative sample of participants in the U.K. Recruitment took place between June 11th and 13th, 2020. This was a period of decline in daily cases following the "first wave", with the U.K. recording between 1,266 and 1,425 new cases of COVID-19 per day during the recruitment period.[5] During this time, the U.K. government was in the process of easing initial "lockdown" restrictions that had been in place since March 26th, 2020.

Method

Participants

A representative sampling procedure was applied via the *Prolific* online platform that stratified across the same age, sex and ethnicity subgroupings used in the most recent U.K. census and sampled participants in each subgroup in proportion to the national population. The final sample consisted of 430 U.K. residents (male = 209, female = 221) aged between 19 and 82 years ($M = 46.1$, $SD = 15.5$). *A priori* power analyses conducted using G*Power (Faul, Erdfelder, Lang, Buchner, 2007), indicated that this sample was sufficient to detect a medium correlation ($r = .30$) with 99% power. Participants were compensated (GBP) £1.90 for their participation.

Materials

---

[4] Pre-registration document can be accessed at this link: https://aspredicted.org/RER_GWJ

[5] The peak of the "second wave" in the U.K. occurred on November 11th, 2020 (22,950 daily cases), and the "third wave" on January 8th, 2021 (68,053 daily cases).

Perceived Threat of COVID-19. Participants' subjective perceptions of the threat posed by COVID-19 was measured using an adapted form of the primary and secondary threat-appraisal measure employed by Tomaka et al. (1993). To measure primary threat appraisal of COVID-19 ("perceived threat"), participants were asked to indicate to what extent they felt COVID-19 was threatening to their health on a sliding scale ranging from: 1 = *Not at all Threatening*, 100 = *Very Threatening*. To measure secondary threat appraisal ("perceived resources") participants were asked to indicate to what extent they personally felt that they had sufficient resources to deal with the threat of COVID-19 to their health (1 = *Not at all sufficient*, 100 = *Completely sufficient*). Consistent with Tomaka et al. (1993), threat appraisal was calculated as an index of participants' perceived threat to their health compared to their perceived resources. However, due to the use of a 100-point scale, threat-appraisal was calculated using the following formula $\frac{Threat}{Threat + Resource}$ in order to avoid issues with variance that would arise from a simple ratio.

Desire for Tighter Norms. Participants' desire for greater cultural tightness was measured using an adapted version of the six-item Tight-Looseness Scale developed by Gelfand et al. (2011). The original Gelfand et al. (2011) scale asked participants to rate how tight they perceived the current norms of their country to be in descriptive terms (e.g. "In this country, there *are very clear* expectations for how people should act in most situations"). These items were reworded in prescriptive terms so as to gauge participants' desire for increased cultural tightness ("e.g. In this country, we *need clearer* expectations for how people should act in most situations"). Participants indicated their level of agreement using a 7-point Likert scale (1 = *Strongly Disagree*, 7 = *Strongly Agree*). Items were averaged such that higher scores indicated greater desire for tight cultural norms (a = .79).

Moralization of COVID-19 Behaviours. Participants' moralization of COVID-19 mitigating behaviours ("COVID-19 behaviours") were measured by presenting participants with a series of six COVID-19 behaviours (e.g., "Someone never hand-washing after being in public places", "Someone leaving the house without wearing a facemask"), and asking them to rate the moral wrongness of each action. Responses were indicated on a 9-point scale (1 = *Not at all morally wrong*, 9 = *Extremely morally wrong*), and averaged such that higher scores indicated greater moralization of COVID-19 behaviours ($a = .80$).

Social and Economic Conservatism. Participants' social and economic conservatism was measured using two single-items questions. For social conservatism, participants were asked to indicate their political beliefs from left/liberal to right/conservative on social issues (e.g., immigration, homosexual marriage, abortion). For economic conservatism, participants were asked to indicate their political beliefs from left/liberal to right/conservative on issues of the economy (e.g., social welfare, government spending, tax cuts). Responses were indicated on a 7-point scale (1 = *Left/liberal* to 7 = *Right/conservative*), such that high scores indicated more conservativism.

Neuroticism. Neuroticism was measured using the relevant subscale of the Big-Five Inventory (Soto & John, 2017). Participants were presented with a number of characteristics (e.g. "Worries a lot"), and asked to indicate the extent to which each applied to them. Responses were indicated on a 5-point scale (1 = *Strongly disagree* to 5 = *Strongly agree*), and averaged such that higher scores indicated higher Neuroticism ($a = .93$).

Procedure

Responses were collected as part of a larger project looking at the effect of individual differences in COVID-19 threat perceptions.[6] Prior to initiating the study, participants were informed that the study was interested in how the COVID-19 pandemic had influenced thinking about social issues, and that they would be asked to answer various questionnaires about their personality, beliefs and attitudes. Once participants completed the survey, they were debriefed and thanked for their participation.

Results

Descriptive statistics and zero-order correlations for all focal variables are presented in Table 1. As predicted, there was a significant positive correlation between participants' appraisals of COVID-19 as threatening to their health, and their desire for cultural tightness in their country, as well as their moralization of COVID-19 behaviours. Further, as predicted, there was a positive relationship between desire for cultural tightness and moralization.

Table 1

Means, standard deviations, and correlations for U.K. participants (N = 430)

Variable	M	SD	1	2	3	4	5	6
1. Threat Appraisal	0.46	0.19						
2. Desire for Tight Norms	4.40	1.16	.20**					
3. COVID-19 Moralization	6.50	1.50	.40**	.27**				
4. Social Conservatism	3.50	1.56	.00	.28**	.08			
5. Economic Conservatism	3.85	1.45	-.04	.18**	.06	.73**		
6. Age	46.13	15.48	.14**	.01	.15**	.21**	.20**	

---

[6] Full survey materials are available at the following link: https://osf.io/yfprs/?view_only=7a8ab67dc0934e64aa78fcd9fc773837

7. Gender (M = 1, F = 2)	.11*	.00	.10*	-.03	-.08	.02

Note. *M* and *SD* are used to represent mean and standard deviation, respectively.
* *p* < .05. ** *p* < .01.

Consistent with past research suggesting high conservatism individuals are more likely to endorse collectivist values (e.g. Graham et al., 2009), we found a positive relationship between both social and economic conservatism and desire for cultural tightness (see Table 1). Desire for cultural tightness was also positively associated with participant age and female gender (see Table 1). Partial correlation analyses were performed to investigate whether the predicted relationships between threat appraisal, desire for cultural tightness, and COVID-19 moralization could be accounted for by these demographic variables. Results revealed no change in significance or direction in the associations between threat appraisal and desire for cultural tightness ($r_{partial}$ = 15, p < .001), moralization of COVID-19 behaviours ($r_{partial}$ = .09, p < .001), or between desire for cultural tightness and moralization ($r_{partial}$ = .26, p < .001) when partialling out variance accounted for by social and economic conservatism, age, and gender.

In order to test the prediction that a desire for cultural tightness would mediate the relationship between COVID-19 threat appraisal and moralization, we conducted a bootstrapped indirect effects analysis with 5,000 bootstrapped samples using the PROCESS package (Hayes, 2013). As predicted, there was a significant indirect effect of COVID-19 threat appraisal on moralization *via* the desire for tighter cultural norms (see Figure 1), *IE* = .30, Boot SE = .11 , p < .001, 95% CI = [.11, .54].

Figure 1

Indirect Effect of COVID-19 Threat Appraisal on Moral Condemnation of COVID-19 Behaviours via Desire for Tighter Norms in Study 1 (U.K.) with 5000 bootstrapped samples
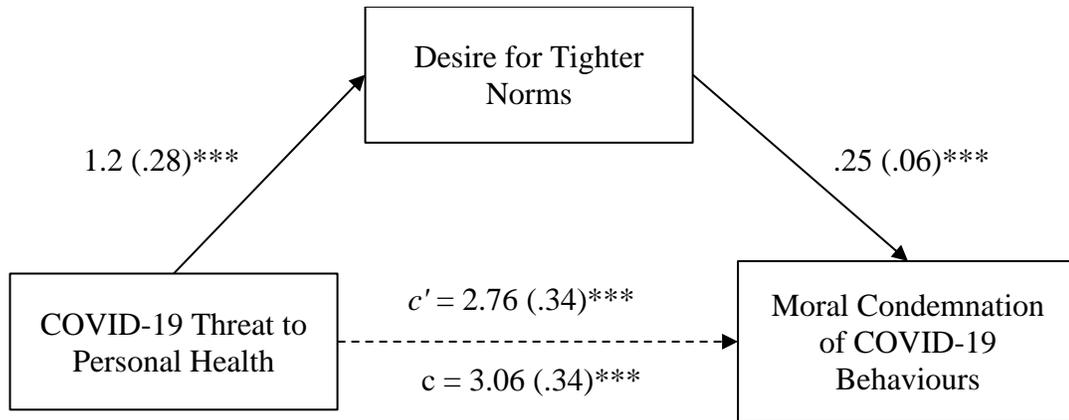
Note. All values represent unstandardized coefficients with standard errors in parentheses

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

The cross-sectional nature of the data made it important to account for the possibility that the observed pattern of results could not be explained by general propensities toward anxiety and threat-sensitivity, as opposed to variance in COVID-19 specific appraisals. Thus, we conducted the same indirect effect analysis after adding individual differences in Big-Five Neuroticism (Soto & John, 2019) as a covariate in the model. To further test the robustness of the findings, social and economic conservatism, age and gender were also included as covariates in the model. Results indicated that adding the covariates did not change the significance or direction of the indirect effect of threat appraisal on moralization of COVID-19 behaviours *via* desire for cultural tightness, *IE* = .34, Boot SE = .11, $p < .001$, 95% CI = [.14, .60].

Study 2

Study 2 sought to replicate the findings of Study 1 in a sample drawn from the U.S. Recruitment took place between June 24th and 26th,[7] 2020. This was a period of increasing daily cases at the beginning of the "second wave", with the U.S. recording between 35,876 and 45,994 new cases of COVID-19 per day during the recruitment period.

Method

Participants

A representative sampling procedure was applied via the *Prolific* online platform that stratified across the same age, sex and ethnicity subgroupings used in the most recent U.S. census, and sampled participants in each subgroup in proportion to the national population. The final sample consisted of 830 U.S. residents (male = 406, female = 417, self-identified = 7) aged between 18 and 82 years ($M = 45.0$, $SD = 16.0$). *A priori* power analyses conducted using G*Power (Faul, Erdfelder, Lang, Buchner, 2007), revealed that this sample was sufficient to detect a medium correlation ($r = .30$) with 99% power.

Materials and Procedure

Subjective COVID-19 threat, desire for tighter norms ($a = . 82$), conservatism, and Big-Five Neuroticism ($a = .93$) were measured using the same items as Study 1.

Moralization of COVID-19 Behaviours. Participants' moralization of COVID-19 behaviours were measured using the measure in Study 1 adapted for the U.S. context.

---

[7] The peak of the "second wave" in the U.S. occurred on July 16th, 2020 (75,820 daily cases), and the "third wave" on January 2nd, 2021 (208,731 daily cases).

Specifically, behaviours that were subject to legal regulation in some states within the U.S. (e.g., mask-wearing) were excluded from the measure ($a = .92$).

Results

Descriptive statistics and zero-order correlations for all focal variables are presented in Table 2. Replicating the findings from Study 1, there was a significant positive correlation between participants' appraisals of COVID-19 as threatening to their health and their desire for cultural tightness in their country, as well as their moralization of COVID-19 behaviours. Further, as predicted, there was a positive relationship between desire for cultural tightness and moralization. As with Study 1, there was a positive relationship between both social and economic conservatism, and desire for cultural tightness (see Table 2).

Partial correlation analyses were conducted to investigate whether the predicted relationships between threat appraisal, desire for cultural tightness and moralization could be accounted for by conservatism or by demographic factors. Replicating Study 1, results revealed no change in significance or direction in the associations between threat appraisal and desire for cultural tightness ($r_{partial} = .11, p < .001$), moralization of COVID-19 behaviours ($r_{partial} = .42, p < .001$), or between desire for cultural tightness and moralization ($r_{partial} = .24, p < .001$) when partialling out variance accounted for by social and economic conservatism, age, and gender.

Table 2

Means, standard deviations, and correlations for U.S. Participants ($N = 830$)

Variable	*M*	*SD*	1	2	3	4	5	6

154

	M	SD					
1. Threat Appraisal	0.48	0.21					
2. Desire for Tight Norms	4.11	1.37	.11**				
3. COVID-19 Moralization	6.40	2.06	.42**	.24**			
4. Social Conservatism	3.19	1.82	-.20**	.20**	-.22**		
5. Economic Conservatism	3.57	1.80	-.26**	.15**	-.28**	.81**	
6. Age	44.99	16.04	.06	-.17**	-.01	.16**	.15**
7. Gender (M = 1, F = 2)			.05	-.08*	.05	-.09**	-.10**	.02

Note. *M* and *SD* are used to represent mean and standard deviation, respectively.
* indicates *p* < .05. ** indicates *p* < .01.

As with Study 1, indirect effects analyses using bootstrapped samples were performed using the PROCESS package (Hayes, 2013) to test the prediction that a desire for cultural tightness would mediate the relationship between COVID-19 threat appraisal and moralization (see Figure 2). Results replicated the pattern of findings in Study 1. Specifically, we found a significant indirect effect of COVID-19 threat appraisal on moralization *via* desire for cultural tightness, *IE* = .21, *Boot SE* = .08, *p* < .001, 95% CI = [.07, .37]. Further, as with Study 1, we found that this indirect effect remains significant when individual differences in neuroticism, social and economic conservatism, age and gender are added as covariates to the model, *IE* = .41, *Boot SE* = .11, 95% CI = [.22, .63].

Figure 2

Indirect Effect of COVID-19 Threat Appraisal on Moral Condemnation of COVID-19 Behaviours via Desire for Tighter Norms in Study 2 (USA) with 5000 bootstrapped samples
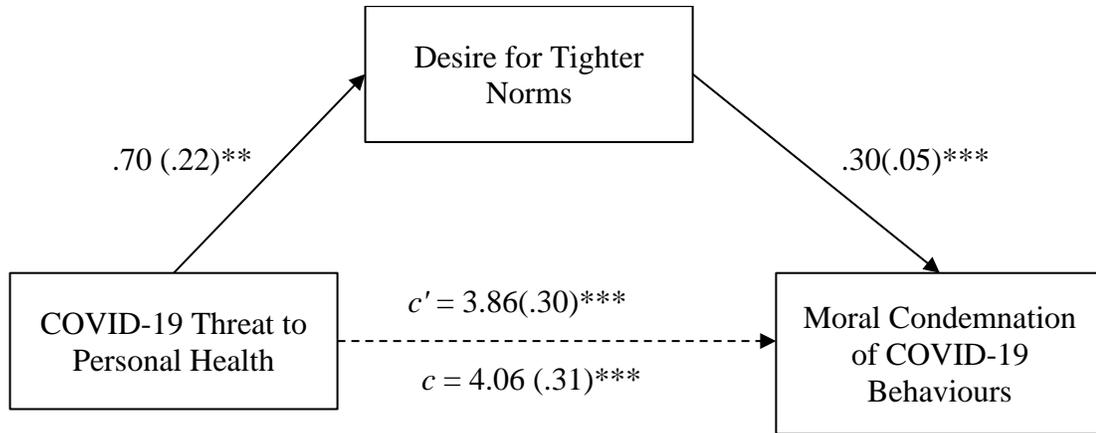
Note. All values represent unstandardized coefficients with standard errors in parentheses
* indicates $p < .05$, ** indicates $p < .01$, ***indicates $p < .001$

General Discussion

The current study aimed to investigate how individual COVID-19 threat appraisals relate to desires for cultural tightness, and moral condemnation of COVID-19 behaviours. We predicted threat-appraisal of COVID-19 would be associated with greater desire for cultural tightness, and greater moral condemnation of noncompliance with COVID-19 behaviours. Further, we predicted that desire for cultural tightness would mediate the relationship between COVID-19 threat-appraisal and moral condemnation of COVID-19 behaviours. Our results supported all three of our hypotheses. More specifically, across representative samples from the U.K. and U.S. that were stratified across the same age, sex and ethnicity subgroupings used in the most recent census, we found individuals who perceived COVID-19 as a threat to their personal health that they did not have sufficient resources to overcome, were more likely to desire a general increase in cultural tightness in their country, and also were more likely to

156

condemn others' failures to engage in COVID-19 mitigating behaviours (e.g. failing to wash their hands). Further, we found an indirect effect of threat-appraisal on moral condemnation of COVID-19 behaviours *via* desire for cultural tightness. These patterns of results remained significant after controlling for age, gender, social and economic conservatism, and individual differences in Big-Five neuroticism.

Threat and Cultural Tightness for the Individual

As far as we are aware, this is the first study to investigate the relationship between COVID-19 threat-appraisals and the desire for cultural tightness at the individual level. Our findings that threat-appraisals of COVID-19 were positively associated with desire for cultural tightness are conceptually consistent with Tight-Loose culture theory, and largely consistent with previous work demonstrating the link between ecological threat and cultural tightness (e.g., Gelfand et al., 2011). By demonstrating the relationship between threat-appraisal and desire for culture tightness at the individual level, our results also corroborate evidence on the relationship between ecological threat and emergence of cultural tightness at multiple levels of analysis such as at the between-country (Gelfand et al., 2011) and within-country levels (Gelfand et al., 2017; Harrington & Gelfand, 2014). In addition, our findings add to current understanding by applying the threat-appraisal approach of Folkman & Lazarus (1991) and (Blascovich & Tomaka, 1996) as a framework for explaining how individuals' appraisals of ecological threats may lead to a desire for cultural tightness. Specifically, we found the proportion of the perceived health threat of COVID-19 and the resources available to deal with the threat was related to the desire for general cultural tightness in participants' own countries.

Moralization in Response to COVID-19 Threat

In addition to desire for cultural tightness, we also found that individual differences in threat-appraisal of COVID-19 were positively associated with moralization of COVID-19 mitigating behaviours (e.g., washing hands, maintaining physical distance). Such findings are consistent with a growing body of research suggesting that meaningful variance can exist between individuals in their tendency to morally condemn COVID-19 behaviours (Bor et al., 2020; Francis & McNabb, 2020; Graso et al., 2021). Our findings are also consistent with recent evidence suggesting that individuals' moralization of COVID-19 may be strongly motivated by desires for self-protection or to advance self-interest (Bor et al., 2020; Francis & McNabb, 2020).

As far as we are aware, this is also the first study to demonstrate the relevance of moralization in the relationship between threat-appraisal and desire for cultural tightness. While there has been previous exploration of how cultural tightness may be related to the perceived moral justifiability of norm violations (Gelfand et al., 2011; Mrazek et al., 2013), such work has been carried out exclusively at the group-level. As such, it has often been suggested that stricter moral condemnations for moral violations may be a consequence or manifestation of cultural tightness (Gelfand et al., 2011; Graham et al., 2016). While our findings cannot offer direct inferences of causal order, they do suggest that higher levels of moral condemnation may accompany desires for cultural tightness in the context of threat-appraisals. This leaves open the possibility that moralization at the individual level may facilitate cultural tightening rather than being a mere reflection of cultural tightness.

It is worth mentioning that although statistically significant, the indirect effect model analyses suggested that desire for cultural tightness did not fully mediate the relationship between threat-appraisal and moralization of COVID-19 behaviours. Such findings suggest that

while there may be overlap between desire for cultural tightness and moralization in response to threat-appraisals, they remain distinguishable constructs.

Finally, our findings have important implications for policy makers considering the long-term implications of threat-based message framing around the COVID-19 (and potential future) pandemic. While cultural tightness may facilitate effective social coordination against ecological threats, it comes with trade-offs related to encouraging group-conflict and harmful obedience, and decreasing social creativity (Gelfand et al., 2020; Jackson et al., 2020). Similarly, while moralization of COVID-19 has been found to increase compliance with government mandates (Francis & McNabb, 2020), it has also been linked with greater acceptance of collateral human costs (e.g. social shaming, police abuses of power) arising from COVID-19 related efforts, compared to non-COVID-19 related efforts. Thus, while threat-based framing may lead to more effective group-level responses to COVID-19 in the short-term, it may have undesired long-term consequences.

Limitations and Future Directions

Although the studies have strengths in terms of ecological validity and generalizability, its cross-sectional nature limits the range of inferences that can be drawn. With specific regard to our mediation analyses, while our pattern of results are consistent with extant theorising on the emergence of cultural tightness (Gelfand et al., 2011; Roos et al., 2015), the cross-sectional design makes it difficult to infer the order of causality between threat-appraisal, desire for cultural tightness and moralization of COVID-19 (Lemmer & Gollwitzer, 2017). In a similar vein, while our results point to potential mechanisms by which threat-appraisals of COVID-19 at the individual level may lead to cultural change, our study was not designed to examine evidence of cultural change. While the aim of our studies were to investigate the potential existence of a

shared function between moralization and cultural tightness in response to COVID-19 threat-appraisals, our results suggest that more detailed investigation of this shared function may be warranted. To this end, an important avenue for future research would be the application of longitudinal methods that are better suited to establishing causal order, and more directly observing cultural change.

Conclusion

The COVID-19 pandemic is a unique generational event affecting virtually every human society on earth, however, its potential impact on human culture remains poorly understood (Kashima et al., 2021). In the face of growing evidence that cultural differences play an important role in the effectiveness of pandemic responses, research into the cultural dynamics arising from the COVID-19 pandemic has never been more timely. By studying these dynamics at the individual level, our study provides important perspectives into the psychological mechanisms that may drive cultural change in response to COVID-19.

References

Blascovich, J. (2008). Challenge, threat, and health. In *Handbook of motivation science* (pp. 481–493). The Guilford Press.

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology*, *28*, 1–51.

Bor, A., Lindholt, M. F., Jørgensen, F. J., & Petersen, M. B. (2020). *Moralizing Physical Distancing during the COVID-19 Pandemic–Personal Motivations Predict Moral Condemnation*.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*(2), 477–496. https://doi.org/10.1037/a0029065

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A., & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology*, *50*(5), 992.

Francis, K., & McNabb, C. (2020). *Moral Decision-Making during COVID-19: Moral judgments, moralisation, and everyday behaviour*. OSF Preprints. https://doi.org/10.31219/osf.io/jvfds

Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., & Hens, N. (2020). Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Eurosurveillance*, *25*(17), 2000257. https://doi.org/doi:https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257

Gelfand, M. J., Caluori, N., Jackson, J. C., & Taylor, M. K. (2020). The cultural evolutionary trade-off of ritualistic synchrony. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1805), 20190432. https://doi.org/10.1098/rstb.2019.0432

Gelfand, M. J., Harrington, J. R., & Jackson, J. C. (2017). The Strength of Social Norms Across Human Groups: *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691617708631

Gelfand, M. J., Jackson, J. C., Pan, X., Nau, D., Pieper, D., Denison, E., Dagher, M., Van Lange, P. A. M., Chiu, C.-Y., & Wang, M. (2021). The relationship between cultural tightness–looseness and COVID-19 cases and deaths: A global analysis. *The Lancet Planetary Health*, *5*(3), e135–e144. https://doi.org/10.1016/S2542-5196(20)30301-6

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Ferrer, M., Fischlmayr, I. C., … Yamaguchi, S. (2011). Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science*, *332*(6033), 1100–1104.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.

Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies. *Current Opinion in Psychology*, *8*, 125–130. https://doi.org/10.1016/j.copsyc.2015.09.007

Graso, M., Chen, F. X., & Reynolds, T. (2021). Moralization of Covid-19 health response: Asymmetry in tolerance for human costs. *Journal of Experimental Social Psychology*, *93*, 104084. https://doi.org/10.1016/j.jesp.2020.104084

Harring, N., Jagers, S. C., & Löfgren, Å. (2021). COVID-19: Large-scale collective action, government intervention, and the importance of trust. *World Development*, *138*, 105236. https://doi.org/10.1016/j.worlddev.2020.105236

Harrington, J. R., & Gelfand, M. J. (2014). Tightness-looseness across the 50 united states. *Proceedings of the National Academy of Sciences*, *111*(22), 7990–7995. https://doi.org/10.1073/pnas.1317937111

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* (pp. xvii, 507). Guilford Press.

Hogarth, R. M. (2005). The challenge of representative design in psychology and economics. *Journal of Economic Methodology*, *12*(2), 253–263. https://doi.org/10.1080/13501780500086172

Jackson, J. C., Gelfand, M., & Ember, C. R. (2020). A global analysis of cultural tightness in non-industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1930), 20201036. https://doi.org/10.1098/rspb.2020.1036

Kashima, Y., Dennis, S., Perfors, A., & Laham, S. M. (2021). Culture and global societal threats: COVID-19 as a pathogen threat to humanity. *Group Processes & Intergroup Relations*, *24*(2), 223–230. https://doi.org/10.1177/1368430220982075

Kemmelmeier, M., & Kühnen, U. (2012). Culture as Process: The Dynamics of Cultural Stability and Change. *Social Psychology*, *43*(4), 171–173. https://doi.org/10.1027/1864-9335/a000117

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

Maaravi, Y., Levy, A., Gur, T., Confino, D., & Segal, S. (2021). "The Tragedy of the Commons": How Individualism and Collectivism Affected the Spread of the COVID-19 Pandemic. *Frontiers in Public Health*, *9*, 37. https://doi.org/10.3389/fpubh.2021.627559

Markus, H. R., & Kitayama, S. (2010). Cultures and Selves: A Cycle of Mutual Constitution. *Perspectives on Psychological Science*, *5*(4), 420–430. https://doi.org/10.1177/1745691610375557

Meijen, C., Turner, M., Jones, M. V., Sheffield, D., & McCarthy, P. (2020). A Theory of Challenge and Threat States in Athletes: A Revised Conceptualization. *Frontiers in Psychology*, *11*, 126. https://doi.org/10.3389/fpsyg.2020.00126

Mrazek, A. J., Chiao, J. Y., Blizinsky, K. D., Lun, J., & Gelfand, M. J. (2013). The role of culture–gene coevolution in morality judgment: Examining the interplay between tightness–looseness and allelic variation of the serotonin transporter gene. *Culture and Brain*, *1*(2–4), 100–117. https://doi.org/10.1007/s40167-013-0009-x

Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *The Journal of Economic Perspectives*, *14*(3), 137–158.

Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals without allies moralize more? *Evolution and Human Behavior*, *34*(2), 78–85. https://doi.org/10.1016/j.evolhumbehav.2012.09.006

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Social and Personality Psychology Compass*, *13*(12), e12511. https://doi.org/10.1111/spc3.12511

Roos, P., Gelfand, M., Nau, D., & Lun, J. (2015). Societal threat and cultural variation in the strength of social norms: An evolutionary basis. *Organizational Behavior and Human Decision Processes*, *129*, 14–23. https://doi.org/10.1016/j.obhdp.2015.01.003

Rozin, P. (1999). The Process of Moralization. *Psychological Science*, *10*(3), 218–221.

WHO Coronavirus (COVID-19) Dashboard. (n.d.). Retrieved September 30, 2021, from

https://covid19.who.int

5.3.   Summary of Chapter 5

In Chapter 5 we found across representative samples in the U.S. and U.K. that threat-appraisal of COVID-19 was positively associated with moralization of others' COVID-19 spreading behaviours. Taken together, the findings of Chapters 3 to 5 provide empirical support for the relevance of threat-appraisal on moralization by showing that: (1) those who are dispositionally more threat-sensitive (and thus more likely to appraise situations as threatening) are also more likely to morally condemn third-party actions; (2) individuals' experience of threat-appraisal emotions in daily life predict their moral judgments; and (3) those who appraise the COVID-19 pandemic as threatening to themselves, are more likely to morally condemn others' engagement in COVID-19 spreading behaviours.

5.3.1.   Relationship between Threat-Appraisal and Moralization not Accounted for by Social Conservativism

Importantly, across both Chapters 3 and 5, we observed that the individual differences in threat-appraisal/moralization relationships remained robust even after controlling for participants' social conservativism. These findings challenge currently influential accounts of the link between threat-appraisal and ideological beliefs, which suggest that those who endorse conservative ideals are particularly motivated to obtain epistemic closure in the face of potential threats (Jost, 2012) and thus are more likely to morally condemn others' behaviours (Janoff-Bulman et al., 2008). Indeed, directly contradicting of such accounts, in Study 2 of Chapter 5, we found among U.S. participants that social conservativism was negatively associated with moralization of COVID-19 spreading behaviours. Such findings support the notion that threat-appraisal/moralization processes reflect more general cognitive and behavioural mechanisms,

rather than reflecting dispositional traits that are only observable in individuals who have a propensity toward certain ideological beliefs.

6

Chapter 6
Moralization of Attitudes following the Release
of Special Prosecutor Robert Mueller's Report
on Collusion between the Trump Campaign and
Russia during the 2016 Presidential Election

Chapter 6 Moralization of Attitudes following the Release of Special Prosecutor Robert Mueller's Report on Collusion between the Trump Campaign and Russia during the 2016 Presidential Election.

6.1. Preface

The aim of Chapter 6 was to investigate whether – as predicted by the action-based model – moralization influenced the degree of certainty that individuals felt about future events. The manuscript presented in this chapter was submitted for publication in *the Journal of Political Psychology* and is presented in author-submitted format.

Across Chapters 3 to 5, we found consistent support for the premise of the action-based moralization model that individuals would be particularly motivated to moralize when they perceive the existence of a potential threat. In Chapter 6, we sought to evaluate another important premise of the action-based moralization model – specifically that moralization is an adaptive functional response to threat because it facilitates future action by reducing epistemic uncertainty. To this end, we conducted a longitudinal study around the release of Special Prosecutor Robert Mueller's report on his Investigations into *Russia's involvement in the 2016 U.S. Presidential Election* (the "*Mueller Report*"). The lead up to the release of the Mueller Report was defined by substantial partisan differences among U.S. residents in perceptions of what the investigation had revealed. The aim of Chapter 6 was to evaluate the potential role that moralization may play in these partisan differences. Specifically, we sought to investigate how moralization might influence epistemic confidence about what the full Mueller Report would reveal in the lead up to its release, and how it would influence subsequent perceptions on what the report had revealed after it was released.

6.2. Manuscript

Moralization of Attitudes following the Release of Special Prosecutor Robert Mueller's Report on Collusion between the Trump Campaign and Russia during the 2016 Presidential Election.

Joshua J. Rhee[1] & Brock Bastian[1]

[1]School of Psychological Sciences, University of Melbourne

Author Note

Correspondence: Joshua J. Rhee, Melbourne School of Psychological Sciences, University of Melbourne, Victoria, 3010, Australia; rheej@student.unimelb.edu.au.

Abstract

Two emergent themes within recent academic and public discussion on affective polarization have been "morality" and "alternative facts. Herein, we sought to investigate the interplay between these two processes within the context of Special Prosecutor Robert Mueller's report on Investigations into Russia's involvement in the 2016 Presidential Election (the "Mueller Report"). A total of 946 US residents provided responses before (T1) and after (T2) the Mueller Report was made public. At T1 participants indicated their level of confidence that the report would contain information demonstrating Russian involvement with President Trump or his campaign during the 2016 election ("T1 Confidence"). At T2, participants indicated whether the Mueller Report had confirmed or contradicted their personal expectations ("T2 Expectation Confirmation") and whether it had strengthened their overall position ("T2 Strengthened Position"). At both timepoints, participants indicated their level of moralization of the Mueller Investigation. Time 1 Moralization was a significant predictor of T2 Expectation Confirmation and T2 Strengthened Position. Finally, T2 Expectation Confirmation was a positive predictor of increase in Moralization from T1 to T2. These findings have important implications for current understanding on how increasing the moral significance of partisan positions may contribute to partisan differences in the evaluation of factual information

Keywords: moralization, partisanship, polarization, epistemics, echo-chambers

Highlights

- Partisan differences in factual beliefs are a well-documented phenomenon in the context of social polarization in the US.

- We conducted a longitudinal study around a highly polarized real-world political event – i.e., the release of the Mueller Report.

- We found that participants' moralization of the investigation predicted whether they perceived the Mueller Report as confirming their expectations and favouring their overall position.

- These findings are likely to be of interest to policy-makers and practitioners interested identifying the psychological processes that exacerbate the formation of polarized factual beliefs.

Interplay between Moralization and Factual Beliefs around the Release of Special Prosecutor Robert Mueller's Report on Russian Involvement during the 2016 Presidential Election

In 2017 special prosecutor Robert Mueller was appointed by the US deputy attorney general to investigate potential links between then-President Donald Trump and operatives in the Russian government who had sought to influence the 2016 US presidential election (the "*Mueller Investigation*"). A striking feature of this investigation as a socio-political phenomenon was the degree of difference between supporters and opposers of President Trump in their perceptions of what the investigation had revealed – prior to any official release of information by the investigation team. For example, a YouGov poll conducted two months prior to any official release about the findings of the investigation, revealed that 71% Clinton voters and 67% of Trump voters reported understanding what Robert Mueller had uncovered to that point either 'fairly well' or 'very well', and further that 70% of Clinton voters believed that the Mueller investigation was likely to conclude that President Trump committed crimes, compared to 5% of Trump voters (YouGov, 2019).

While the existence of partisan differences in factual beliefs is a well-documented phenomenon in the context of social polarization in the US (e.g. Doherty et al., 2019), the circumstances surrounding the Mueller Investigation were relatively unique in that there was an authoritative event – i.e. the release of the official report outlining the findings of the Mueller Investigation (the "*Mueller Report*") – that could confirm or contradict some of the relevant polarized factual beliefs. Such a context provided an opportunity to investigate how people in a highly polarized environment would respond to new belief confirming (or disconfirming)

information. Specifically, we investigated the potential interplay between the construct of moralization – that is, the degree to which an individual attaches moral significance to issues, actions or entities is referred to as moralization (Rhee et al., 2019; Rozin, 1999) – and individuals' updating of their factual beliefs following the release of the Mueller Report.

Moralization, Factual Belief and Polarization

The tendency for partisan opponents to increasingly frame disagreement in terms of "good" versus "evil"; and the tendency for each side to seemingly rely on different sets of facts – have become increasingly important themes of modern affective polarization in US politics. A recent Pew Research survey of 9,895 respondents found that 55% of Republicans and 47% of Democrats viewed members of the opposing party as more immoral than the average American, while also finding that 73% of Americans believed that Democrats and Republicans could not agree on basic facts when it came to important issues facing their country. Despite the co-occurrence of these two phenomena, there has been little empirical research investigating the interplay between moralization and factual beliefs in the context of polarizing real-world political events (Alesina et al., 2020). In this study, we propose that polarization of factual beliefs and moralization are interconnected processes. Specifically, drawing on lay epistemics theory (Kruglanski et al., 2006), we propose that moralization may lead to an increased tendency among individuals to use their existing moralized beliefs to evaluate the epistemic content of new information.

Moralization and Epistemic Certainty

Humans have a fundamental motivation to feel that we have been effective in understanding the truth (Higgins, 2013). In other words, we are motivated to feel effective in

174

establishing what is occurring in our environment, to ensure that our motivations to approach rewards or avoid pain is directed appropriately (Cornwell et al., 2014). According to lay epistemics theory (Kruglanski, 1990) people seek to establish this knowledge about their environment through a process of lay hypothesis generation and evidence gathering (Kruglanski et al., 2010). The nature of the information that is treated as evidence in this process is strongly determined by the underlying motivation of the individual (Kruglanski et al., 2010). For example, when an individual is highly motivated to gain an accurate worldview, they may continually seek out new domain-relevant information in order to ensure that their knowledge is representative of the environment (Gershman, 2019; Kruglanski & Webster, 1996). Alternatively, individuals can also be motivated by the need for non-specific epistemic closure – that is, a need to obtain any definitive answer to some topic in order to minimise confusion and ambiguity (Kruglanski et al., 2010; Kruglanski et al., 1993).

Recent research suggests that one strategy individuals draw on to reduce uncertainty and satisfy their motivation for non-specific closure, may be to use their moral beliefs to inform their process of lay hypothesis generation and evidence gathering. For example, Stanley et al. (2020) found that individuals with high trait motivation for cognitive closure were more likely to attribute objectivity to their moral beliefs. In other words, individuals motivated to see the world as structured, ordered and predictable were willing to accept their moralized beliefs as factual reflections of the world around them. More notably, Liu & Ditto (2013) found that participants who perceived an act as morally wrong (e.g. whether forceful interrogation is wrong) were also more likely to believe in factual claims supporting the low utilitarian value of those acts (e.g. that forceful interrogation does not produce valid intelligence). Importantly, it was also found that participants who reported greater moralization of their positions on each issue, demonstrated a

stronger relationship between seeing the acts as morally distasteful and believing evidence of their low utilitarian value.

These findings suggest that the level of moralization of a belief may influence whether it is used as epistemic evidence for evaluating other factual information. In other words, when beliefs are highly moralized, individuals may use consistency with those beliefs as evidence for how new information they encounter should be evaluated. Further, given the ability of moralized beliefs to provide a sense order and structure (Dechesne & Kruglanski, 2004; Federico et al., 2016; van den Bos et al., 2005), individuals who evaluate new information in this way may feel a greater sense of confidence about the knowability of new information they encounter in the future (van Prooijen & Krouwel, 2019).

Applied to the context of the current study, it is possible that polarizing issues where individuals hold highly moralized positions may be particularly conducive of factual belief polarization, as they increase the likelihood that individuals use consistency with their moralized beliefs as a form of evidence to evaluate new information they encounter. Because moralized beliefs are highly internalised, using such beliefs to evaluate new information would lead to the formation of insular factual beliefs that are not readily updated by objective external evidence.

Compensatory Moralization in Response to Epistemic Uncertainty?

If individuals do use their moralized beliefs to evaluate new information, a salient question is how individuals respond when they encounter incontrovertible information that discredits their moralized beliefs? When an individual encounters credible and substantive information that is in direct conflict with an existing cognition, they are theoretically faced with two options: (1) dispose of the existing cognition, and adopt new cognitions that are consistent with the information; or (2) change the nature of the of the existing cognition such that the

information is no longer inconsistent, or trivialises the challenge presented by the contradictory information (Proulx & Heine, 2006). Although both of these options have their costs and benefits, research has overwhelmingly demonstrated that, people will often tend to favour the option of maintaining their existing cognitive schemas (Hart & Nisbet, 2012; Proulx & Heine, 2006; Swire et al., 2017) through motivated reasoning (Kunda, 1990). Indeed, people presented with compelling evidence that is in direct contravention with their beliefs have often been found to display a "boomerang" or "backfire" effect whereby they actually strengthen the attitudes threatened by the evidence (Byrne & Hart, 2009; Hart & Nisbet, 2012; Trevors et al.,2016).

Perhaps the most influential demonstration of this phenomenon was a longitudinal field-study by Festinger et al. (1956) who followed members of a religious group that predicted the end of world was imminent. In a book entitled "*When Prophecy Fails*", Festinger et al. (1956) presented their findings from interviews conducted before and after the predicted "doomsday" had failed to transpire. Festinger et al. (1956) remarkably observed that in the face of incontrovertible evidence of the falsity of their beliefs, the members of the religious group held onto their beliefs, and indeed, began proselytising them.

Applied to the current inquiry, this past work suggests that, when presented with external evidence that directly contradicts beliefs that an individual is highly committed to, they will most likely respond by looking for alternative epistemic strategies that allow them to maintain their beliefs. Starting from the premise that individuals who are highly motivated by their need for cognitive closure are more likely to adopt their moral beliefs as epistemic evidence (Stanley et al., 2020), it is possible that individuals may use moralization as a compensatory strategy for maintaining their current beliefs, when presented with belief-disconfirming information. In other words, while people who have moralized beliefs on a topic may have a general propensity to

interpret new information as confirming their moralized positions, when presented with information that strongly challenges their position, they may also seek to double down by increasing their level of moralization, and thus decreasing the need to take external information into account when assessing their beliefs.

The Current Study

Whether individuals use their moralized beliefs as evidence for epistemic evaluation of new information is an important question in understanding the dynamics of social polarization. Given the subjective nature of moralized beliefs, their use as evidence in epistemic evaluation processes provides a powerful explanation for the rise fundamentally different perceptions of reality along partisan lines in US society (Alesina et al., 2020). Despite the importance of better understanding the relationship between moralization and epistemic evaluation in the context of US partisanship, little work to date has investigated these dynamics in the context of real-world political events. To address this gap in understanding, the current study used the context of Special Prosecutor Robert Mueller's Report on Russian involvement during the 2016 Presidential Election to investigate the dynamics between moralization and evaluation of factual information.

The current study aimed to investigate how moralization of beliefs prior to the release of the Full Mueller Report interacted with participants' expectations regarding what the report would find, and their evaluation of the factual contents of the Full Mueller Report after it was released. To this end, we conducted a pre-post study where participants were recruited immediately prior to the scheduled release of the Full Mueller Report, and then subsequently invited to participants in the second part of the study around a week after the report was publicly released.

Figure 1

Conceptual diagram of key variables and hypotheses

Time 1	Report Release	Time 2

Confidence in Predictions of the Mueller Report (**T1 Confidence**)

Moralization (**T1 Moralization**)

Perceived Confirmation of Expectations (**T2 Expectation Confirmation**)

Perceived Support for Previous Position (**T2 Strengthened Position**)

Change in Moralization (**T2 Moralization** controlling for **T1 Moralization**
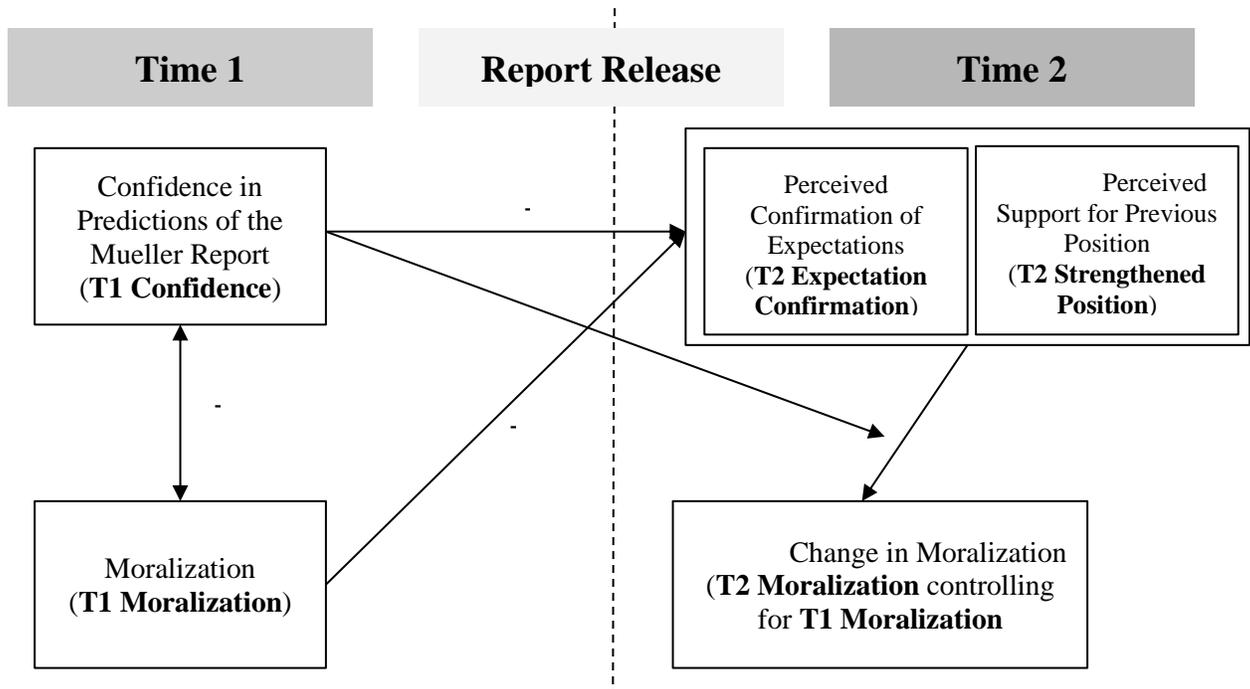
Based on the notion that individuals would use their moralized beliefs to evaluate factual information, we predicted that T1 Moralization would be associated with T1 Confidence about what the report would reveal. We also predicted that T1 Moralization would predict participants' T2 appraisal that the Full Report had confirmed their previous expectation (T2 Expectation Confirmation), and that it had strengthened their overall position (T2 Strengthened Position). Finally, drawing on the notion that individuals will respond to external evidence that directly contradicts their committed beliefs by adopting epistemic strategies that enable belief maintenance, we predicted a significant *T1 Confidence* x *T2 Strengthened Position* interaction effect on change in Moralization between T1 and T2. More specifically, we predicted that those who had a strong commitment to their beliefs about what the Mueller Investigation (T1 Confidence) would be most likely appeal to moralization (T2 Moralization) when the real findings of the Mueller Report were perceived to weaken their previous position (T2

179

Strengthened Position). A summary of these key variables and their expected relationships are

presented in Figure 1.

<center>**Method**</center>

Participants and Procedure

We employed a pre-post design. Time 1 data was collected in the week before the release of the Full Mueller Report, and Time 2 data was collected two weeks after the release of the Full Mueller Report (see Figure 2).

Figure 2

Timeline of Study and major events around the public release of the Special Counsel Robert Mueller's Report on The Investigation into Russian Interference in The 2016 Presidential Election ("Mueller Report")
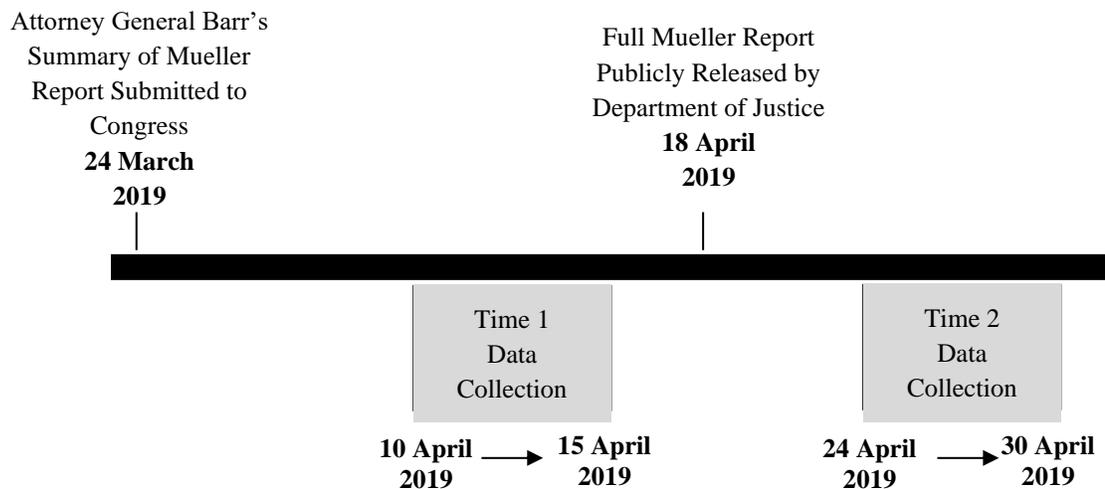
Participants were recruited online via Amazon's Mechanical Turk (MTurk) platform. Participants were required to be above 18 years of age, and a current U.S. resident in order to take part in the study. A total of 1,461 participants took part in the Time 1 data collection that took place from April 10th, 2019 to April 15th, 2019. Two-weeks' after the public release of the Mueller Report by the U.S. Department of Justice participants were invited to complete the Time-2 follow-up survey ($N = 946$). The Time 2 response rate of 64.7% was somewhat higher

than similar studies conducted on the MTurk platform around real-world U.S. political events, which often fall below 60% (Brandt et al., 2015; Turner-Zwinkels et al., 2015). Participants were compensated at a rate of US$0.50 (fifty cents) for each wave of data collection they participated in. The final sample of 946 participants (*Female* = 516, *Male* = 428, *Unspecified* = 2) ranged from 18 to 87 years of age (*M* = 41.5, *SD* = 13.1). Politically, the sample leaned slightly towards supporters of the Democratic party, with 467 voting Democrat, 235 voting Republican, 143 voting for a non-major party and 101 not voting in the 2016 Presidential election.

Materials

Moralization. Participants' level of moralization regarding their positions on "uncovering the truth regarding President Trump's potential involvement with Russia or obstruction of justice" were measured using the 4-item moral conviction scale (e.g. a reflection of your core moral beliefs and convictions?). Responses were recorded on a scale from 1 (Not at all) to 5 (Very Much) and averaged such that higher scores indicated higher moralization.

Time 1 Measures

Time 1 Confidence. Participants' confidence regarding the contents of the Full Mueller Report were assessed on a 100-point confidence slider (0% to 100%). For all questions, the starting point of the slider was at the midpoint of the scale. More specifically, participants indicated the degree to which they were confident that the Full Report prepared by Special Prosecutor Robert Mueller will: (1) "Contain information suggesting that President Trump engaged in obstruction of Justice"; and (2) "Contain information suggesting that President Trump or someone from his campaign worked with Russia to influence the 2016 election". To account for the polarizing nature of the topic, we conceptualised high confidence as use of extreme values on the scale. To this end, responses on this scale were folded along the midpoint

and averaged ($r = .62$) such that an initial rating of "50" would indicate the lowest level of confidence, and ratings of "0" or "100" would indicate the highest level of confidence.

Time 1 Trust in Report. Participants' level of confidence in the accuracy of the Mueller Report itself was measured using a one-item question ("Contain a fair and accurate assessment of the facts surrounding President Trump's involvement with Russia and/or obstruction of justice."). Participants responded to the item using a 100-point confidence slider (0% to 100%). The starting point of the slider was at the midpoint of the scale.

Time 1 Trump Approval. To control for participants' personal support of President Trump, participants were asked a one item question: "Overall, do you approve or disapprove of the way Donald Trump is handling his job as President?". Participants responded to this item on a 6-point scale (1= Strongly approve; 6 = Strongly disapprove).

Time 2 Measures

Time 2 Expectation Confirmation. Participants' perceptions of whether the released Full Muller Report had confirmed their previous expectations was measured using a two-item measure. Participants were asked to indicate on a 100-point slider (-50 = Completely Contradicted to 50 = Completely Confirmed) whether the Full Mueller Report had confirmed their personal expectations regarding whether President Trump or someone from his campaign: (1) worked with Russia to influence the 2016 election; and (2) engaged in obstruction of Justice. Responses were averaged such that higher scores represented strong expectation confirmation ($r = .80$).

Time 2 Strengthened Position. Participants' perceptions that the contents of the Full Mueller Report supported their overall position was measured using one-item: "do you think that the contents of the Full Report to strengthened or weakened your position on President Trump".

Participants responded on a 7-point scale (1 = Weaken a lot; 4 = about the same; 7 = Strengthen

a lot)

Table 1

Means, standard deviations, and correlations

Variable	M	SD	1	2	3	4	5	6	7	8
1. Age	39.50	13.10								
2. Gender	1.46	0.50	.16**							
3. T1 Trump Approval	4.60	1.87	-.16**	.08*						
4. T1 Confidence	33.58	14.42	.23**	.02	-.15**					
5. T1 Trust in Investigation	32.47	16.52	.07*	-.04	.16**	.49**				
6. T1 Moralization	3.42	1.26	.12**	.08*	.34**	.13**	.24**			
7. T2 Expectation Confirmation	10.75	33.68	-.07*	.12**	.45**	.03	.12**	.26**		
8. T2 Strengthened Position	5.11	1.50	.17**	.07*	-.05	.29**	.19**	.23**	.12**	
9. T2 Moralization	3.35	1.27	.11**	.10**	.35**	.10**	.21**	.69**	.28**	.23**

Note. M and *SD* are used to represent mean and standard deviation, respectively.
* indicates $p < .05$. ** indicates $p < .01$

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

Statistical analysis

Descriptive statistics and zero-order correlations for all relevant Time 1 and Time 2 variables are presented in Table 1.

To test our main hypotheses, we conducted our analyses in R using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) to fit linear mixed-effects models, and lmerTest (Kuznetsova, Brockhoff, & Cristensen, 2017) to compute p-values. To aid interpretability and reduce convergence issues, all variables were standardized and scaled. Moreover, all time-level variables were Party-mean centered (i.e. the Political Party means was subtracted from each score) and allowed to vary randomly across participants at the person-level (i.e. an intercept-only model).

Time 1 Moralization predicting Time 2 Expectation Confirmation.

To test the first hypothesis that T1 Moralization would predict T2 Expectation Confirmation and T2 Strengthened Position, we ran a mixed-effects model with Participants' Political Party ($N = 8$) as a random effect. Support for President Trump (T1 Trump Approval) was included as a covariate to control for baseline differences in moralization of the Mueller Investigation based on participants' existing views of President Trump. In order to control for cynical responses from participants who might indicate that the report confirmed their expectations by being inaccurate, we included T1 Trust in the Report as a covariate in the model.

As predicted, we found that Time 1 Moral Conviction was a significant predictor of Time 2 perceptions that the Mueller Report had confirmed participants' previous expectations (T2 Expectation Confirmation), and that it had strengthened their overall position (T2 Strengthened Position) on President Trump (see Table 2).

Table 2

Linear Mixed effects models predicting Time 2 perceptions of whether the released Full Muller Report had confirmed participants' previous expectations (Expectation Confirmation) and supported their overall position (Strengthened Position)

Predictors	Time 2 Expectation Confirmation			Time 2 Strengthened Position		
	Estimates	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.28	0.04 – 0.53	**.023**	0.93	0.88 – 0.98	**<.001**
T1 Confidence	0.09	0.03 – 0.16	**.004**	0.06	0.04 – 0.08	**<.001**
T1 Moral Conviction	0.09	0.04 – 0.15	**.001**	0.05	0.04 – 0.07	**<.001**
T1 Trump Approval	0.22	0.16 – 0.27	**<.001**	-0.01	-0.03 – 0.00	.115
T1 Trust in Investigation	-0.01	-0.08 – 0.05	.653	0.01	-0.01 – 0.03	.166
Random Effects						
σ^2	0.71			0.07		
$\tau_{00 \text{ Party\_2016}}$	0.09			0.00		
ICC	0.11			0.04		
N $_{\text{Party\_2016}}$	7			7		
Observations	942			943		
Marginal R^2 / Conditional R^2	.08 / .19			.11 / .15		

6.3.   Changes in Moralization Between Time 1 and Time 2

To test the hypothesis that there would be a two-way *T1 Confidence* x *T2 Strengthened Position* interaction on T1 to T2 change in Moralization, we ran a two-level mixed-effects model in which persons ($N = 946$) were nested within Party ($N = 8$). To control for the possibility that

variance in T2 Strengthened Position could be accounted for by participants' political leaning,

T1 Trump Approval was added as a covariate in the model. Contrary to expectations, we found

no significant two-way *T1 Confidence* x *T2 Strengthened Position* interaction effect predicting

T2 Moralization (see Table 3). However, we did find that T2 Strengthened Position was a

significant positive predictor of T2 Moralization, suggesting that participants who saw the Full

Mueller Report as strengthening their overall position on President Trump were also more likely

to report an increase in moralization from Time 1 to Time 2 (see Table 3).

Table 3
Linear Mixed effects models predicting Change in Moralization from Time 1 to Time 2

	Change in Moralization		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.87	0.80 – 0.95	**<.001**
Time	-0.00	-0.02 – 0.01	.638
T2 Strengthen Position	0.06	0.04 – 0.08	**<.001**
T1 Confidence	0.02	0.00 – 0.06	**.043**
T2 Strengthen Position x T1 Confidence	0.01	-0.01 – 0.03	.380
T1 Trump Approval	0.06	0.05 – 0.08	**<.001**
T1 Trust in Investigation	0.03	0.01 – 0.06	**.001**
Random Effects			
σ^2	0.04		
τ_{00} id:Party_2016	0.06		
τ_{00} Party_2016	0.01		
ICC	0.64		
N id	943		
N Party_2016	7		
Observations	1886		

Marginal R^2 / Conditional R^2	0.11 / 0.68

Discussion

The aim of the current study was to investigate the longitudinal interplay between moralization and epistemic confidence in the context of a highly polarized real-world event. Specifically, we looked at how moralization related to epistemic evaluation in the context of Special Counsel Robert Mueller's Report on the Investigation into Russian Interference in the 2016 U.S. Presidential Election. We predicted that those with already moralized positions on the investigation prior to the public release of the Mueller report, would be more likely to perceive it as confirming their expectations and strengthening their overall position when it was released. We further predicted that those who were highly confident about what the report would reveal at Time 1, would report a greater increase in Moralization from Time 1 to Time 2, if they subsequently perceived the report as weakening their overall position on President Trump.

Moralization and Expectation Confirmation

Consistent with our prediction, we found that T1 Moralization predicted participants' T2 appraisal that the Full Report had confirmed their previous expectation (T2 Expectation Confirmation), and that it had strengthened their overall position (T2 Strengthened Position). Such findings suggest that those who had more strongly moralized their positions prior to the Full Mueller Report's release were more likely to believe that the real report had supported their prior beliefs.

These findings build on a growing body of research investigating the degree to which moralized beliefs are treated as forms of objective knowledge (Goodwin & Darley, 2008, 2012; Liu & Ditto, 2013; Stanley et al., 2020). In particular, our findings corroborate the findings of

Liu & Ditto (2013) that the moralization of certain beliefs can increase an individual's willingness to believe the truth of information that supports the value of those beliefs. Indeed, our findings extend the work of Liu & Ditto (2013) by showing that moralized beliefs don't simply lead to an increased willingness to believe in belief-supporting information, it may also increase the propensity to view new information as belief-supporting. This conceptualisation of moralization as an important process influencing epistemic evaluation of new information provides a strong psychological process account for how increasing the moral significance of partisan positions may contribute to the formation of partisan differences in factual beliefs.

Moralization and Belief Disconfirmation

The Mueller Investigation provided a relatively unique opportunity to investigate the impact of an event (the release of the Mueller Report) that could confirm or contradict highly polarized factual beliefs. In the current study, we sought to build on previous findings within the context of cognitive dissonance (Festinger, 1957) and meaning maintenance frameworks (Hart & Nisbet, 2012; Proulx & Heine, 2006; Swire et al., 2017) that individuals who are highly committed to certain beliefs will adopt motivated reasoning strategies that allow them to maintain their existing belief systems. Specifically, we predicted that individuals who were highly committed to their beliefs, would seek to maintain those beliefs in the face of strong contradictory evidence by elevating their moral status.

Contrary to our expectations, we did not find support for the hypothesis that those who had a strong commitment to their beliefs about what the Mueller Investigation would be most likely appeal to moralization when the real findings of the Mueller Report were perceived to weaken their previous position. Indeed, our data suggested that individuals were more likely to

double down in moralizing their beliefs, when they perceive new information as strongly confirming their overall position.

One potential explanation for these findings is that the Full Mueller Report was sufficiently neutral in its conclusions, such that they were unlikely to be interpreted as disconfirming of any particular position (Law, 2019). Alternatively, it is also possible that we failed to observe moralization in response to perceived belief disconfirmation because, by the time we were able to measure changes in moralization, participants had already engaged in this compensatory moralization process. This would mean that they may have already reappraised the contents of the Mueller Report as strengthening their overall position. A similar phenomenon was observed by Festinger (1956) who noted that when believers of an end of world cult were faced with the reality that the world had not ended on the predicted date, they reinterpreted the events as supporting their belief framework. Specifically, the believers suggested that they had, through their faith, averted the end of the world by pleasing the forces behind the predicted apocalypse (Festinger, 1956).

Limitations and Future Directions

The inferences made from this study must be considered in light of some limitations in its methodology. First, while conducting a pre-post study in the context of a real-world political event has the advantage of high ecological validity, it also increases uncertainty about the generalisability of the findings beyond the specific event that we studied. Given the novelty of our findings regarding the relationship between moralization and epistemic evaluation of information, replicating this pattern of findings across a wide range of contexts will be an important direction for future research. Second, collection of survey data at only two time points meant that our study had relatively low temporal resolution, which made it difficult to evaluate

the causal order of certain patterns we observed in our results. Specifically (as noted above) we found that Time 2 perceptions that the Mueller Report strengthened participants' overall positions was positively related to changes in Moralization from Time 1 to Time 2. However, because our Time 2 measure was taken around two-weeks after the release of the Mueller Report, it was difficult to make strong inferences about whether this observed relationship was the product of participants having already engaged in the predicted compensatory moralization process. Another important direction for future research would therefore be to observe the relationship between moralization and belief disconfirming information using methods with higher temporal resolution.

Conclusion

The emergence of partisan divides in factual beliefs has been an insidious feature of modern social polarization that presents substantial obstacles to ongoing social progress. Our findings provide a strong psychological process account of how increasing the moral significance of partisan positions may contribute to partisan differences in the evaluation of factual information. These findings have important implications for researchers and policy-makers seeking to develop interventions aimed at reducing partisan differences in factual beliefs.

References

Alesina, A., Miano, A., & Stantcheva, S. (2020, May). The polarization of reality. In *AEA Papers and Proceedings* (Vol. 110, pp. 324-28).

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 U.S. Presidential Election Campaign. *Journal of Social and Political Psychology*, *3*(2), 211–237. https://doi.org/10.5964/jspp.v3i2.434

Byrne, S., & Hart, P. S. (2009). The boomerang effect a synthesis of findings and a preliminary theoretical framework. *Annals of the International Communication Association*, *33*(1), 3-37.

Cornwell, J. F., Franks, B., & Higgins, E. T. (2014). Truth, control, and value motivations: the "what,""how," and "why" of approach and avoidance. *Frontiers in systems neuroscience*, *8*, 194.

Dechesne, M., & Kruglanski, A. W. (2004). Terror's Epistemic Consequences: Existential Threat and the Quest for Certainty and Closure. In *Handbook of Experimental Existential Psychology* (pp. 247–262). Guilford Press.

Doherty, C., Kiley, J., & Asheer, N. (2019). Partisan antipathy: More intense, more personal. *Pew Research Center*.

Federico, C. M., Ekstrom, P., Tagar, M. R., & Williams, A. L. (2016). Epistemic Motivation and the Structure of Moral Intuition: Dispositional Need for Closure as a Predictor of Individualizing and Binding Morality. *European Journal of Personality*, *30*(3), 227–239. https://doi.org/10.1002/per.2055

Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails.* University of Minnesota Press. https://doi.org/10.1037/10030-000

Gershman, S. J. (2019). Uncertainty and exploration. *Decision*, *6*(3), 277.

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition*, *106*(3), 1339–1366. https://doi.org/10.1016/j.cognition.2007.06.007

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology*, *48*(1), 250–256. https://doi.org/10.1016/j.jesp.2011.08.006

Grgic, G. (2019, March 25). Mueller Report: The Good, the Bad and the Ugly - AIIA. *Australian Institute of International Affairs*. https://www.internationalaffairs.org.au/australianoutlook/mueller-report-good-bad-ugly/

Hart, P. S., & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication research*, *39*(6), 701-723.

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and social psychology review*, *10*(2), 88-110.

Higgins, E. T. (2013). Truth motivation. In K. D. Markman, T. Proulx, & M. J. Lindberg (Eds.), *The psychology of meaning* (pp. 91–114). American Psychological Association. https://doi.org/10.1037/14040-005

Kruglanski, A. W. (1990). Lay Epistemic Theory in Social-Cognitive Psychology. *Psychological Inquiry*, *1*(3), 181–197.

Kruglanski, A. W., Orehek, E., Dechesne, M., & Pierro, A. (2010). Lay Epistemic Theory: The Motivational, Cognitive, and Social Aspects of Knowledge Formation. *Social and Personality Psychology Compass*, *4*(10), 939–950. https://doi.org/10.1111/j.1751-

9004.2010.00308.x

Kruglanski, A. W., Pierro, A., Mannetti, L., & De Grada, E. (2006). Groups as epistemic

providers: Need for closure and the unfolding of group-centrism. *Psychological Review*,

113(1), 84–100. https://doi.org/10.1037/0033-295X.113.1.84

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind:" Seizing" and"

freezing.". *Psychological review*, *103*(2), 263.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480.

Law, T. (2019, April 19). *Here Are the Biggest Takeaways From the Mueller Report*. Time.

https://time.com/5567077/mueller-report-release/

Liu, B. S., & Ditto, P. H. (2013). What Dilemma? Moral Evaluation Shapes Factual Belief.

Social Psychological and Personality Science, *4*(3), 316–323.

https://doi.org/10.1177/1948550612456045

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review

of current definitions, methods, and evidence in moralization research. *Social and

Personality Psychology Compass*, *13*(12), e12511.

Rozin, P. (1999). The process of moralization. *Psychological science*, *10*(3), 218-221.

Skitka, L. J., & Morgan, G. S. (2014). The Social and Political Implications of Moral Conviction.

Political Psychology, *35*, 95–110. JSTOR.

Stanley, M. L., Marsh, E. J., & Kay, A. C. (2020). Structure-seeking as a psychological

antecedent of beliefs about morality. *Journal of Experimental Psychology: General*,

149(10), 1908–1918. https://doi.org/10.1037/xge0000752

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting

inaccurate information. *Journal of experimental psychology: learning, memory, and*

cognition, *43*(12), 1948.

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Winne, P. H. (2016). Identity and
epistemic emotions during knowledge revision: A potential account for the backfire
effect. *Discourse Processes*, *53*(5-6), 339-370.

Turner-Zwinkels, F., van Zomeren, M., & Postmes, T. (2015). Politicization During the 2012
U.S. Presidential Elections: Bridging the Personal and the Political Through an Identity
Content Approach. *Personality and Social Psychology Bulletin*, *41*(3), 433–445.
https://doi.org/10.1177/0146167215569494

van den Bos, K., Poortvliet, P. M., Maas, M., Miedema, J., & van den Ham, E.-J. (2005). An
enquiry concerning the principles of cultural norms and values: The impact of uncertainty
and mortality salience on reactions to violations and bolstering of cultural worldviews.
Journal of Experimental Social Psychology, *41*(2), 91–113.
https://doi.org/10.1016/j.jesp.2004.06.001

van Prooijen, J.-W., & Krouwel, A. P. M. (2019). Psychological Features of Extreme Political
Ideologies. *Current Directions in Psychological Science*, *28*(2), 159–163.
https://doi.org/10.1177/0963721418817755

7

Chapter 7
General Discussion

Chapter 7 General Discussion

The aim of this thesis was to better understand the functions served by the process of moralization for the individual. Specifically, my doctoral research program was developed with the aim of understanding moralization's potential function as a psychological response to threat. To this end, a series of empirical studies were conducted in this thesis that sought to address two broad research questions: (1) can moralization be reliably observed as co-occurring with high threat-appraisal? (2) What are the potential functional benefits of moralization in the context of high threat-appraisal? This final chapter will review the findings from each of the empirical chapters in terms of how they address the aforementioned research questions, before turning to a discussion on the theoretical implications of such findings, limitations, and directions for future research.

7.1. Review of Key Empirical Findings

7.1.1. Can moralization be reliably observed as co-occurring with threat-appraisal?

Chapter 3 applied an individual differences approach to investigate whether general variance between individuals in their propensity to moralize third-party actions reliability coincided with individual differences in threat-sensitivity. In Studies 1 & 2 of Chapter 3, participants were first measured on two converging measures of threat-sensitivity: Behavioural Inhibition Sensitivity (BIS: Carver & White, 1994); and Sensation-Seeking (Zuckerman, 1994). While BIS and Sensation-Seeking were initially conceptualized as separate constructs, measures of both have previously been demonstrated to correlate with behavioural and biological markers of threat-sensitivity such as startle-eyeblink responses (Caseras et al., 2006; Lissek & Powers, 2003), and amygdala activation in response to fearful stimuli (Cools et al., 2005; Zheng et al., 2011). Participants' propensity to moralize concrete actions (Rhee et al., 2019) were then

measured using the Moral Foundations Vignettes ("MFV") vignettes developed by Clifford et al. (2015) - a set of ninety vignettes portraying intentional actions by a third-party that represent violations of a wide range of morally relevant domains.

Across Studies 1 & 2, we found that participants' tendency to moralize third-party actions was significantly associated with both BIS (positively) and Sensation-Seeking (negatively). In Study 2, we found that these relationships remained significant after controlling for individual differences in empathic concern and perspective taking, suggesting that the relationship between threat-sensitivity and moralization may be accounted for by self-oriented (rather than other-oriented) factors. Finally, in Study 3 of Chapter 3, we aimed to more closely examine the nature of the relationship between threat-sensitivity and moralization by asking participants to indicate how they felt about each of the vignettes they rated.

It is important to note that, by definition, moralization describes changes occurring within individuals (see Chapter 1). Thus, longitudinal or within-person studies are widely considered the gold-standard of evidence in moralization research (Brandt et al., 2015; Feinberg et al., 2019; Wisneski et al., 2020). In Chapter 4 we therefore sought to investigate whether the relationship between threat-sensitivity and tendency to moralize that we observed at the between-person level in Chapter 3, could also be observed within-persons. To this end, we employed an ecological momentary assessment approach using the SEMA3 platform (Koval et al., 2019), which used participants' smartphones to administer six surveys spaced out over an interval of two hours per day, over a period of seven days. Each survey first measured participants' current experience of a range of emotions, with the most relevant measured emotions being the threat-appraisal emotions used in Chapter 4, Study 3. Importantly, at each timepoint, participants were asked to rate the moral wrongness of a randomly selected vignette *after* reporting their current emotional

experience, thus allowing us to infer causal order. In order to maximise the ecological validity of the study, participants were presented with moral vignettes that were based on recollections of real-world events (Kruepke et al., 2018). We found that threat-appraisal emotions were a significant predictor of changes in participants' moral judgments. Specifically, when participants experienced greater than average levels of threat-appraisal emotions, they were more likely to judge vignettes as more morally wrong compared to the previous time point.

Chapters 3 & 4 provided evidence to support the existence of a general threat-appraisal/moralization relationship, both in terms of individual differences in threat-sensitivity and within-person changes in experience of threat-appraisal emotions. In Chapter 5, we sought to further test the robustness of this relationship by investigating whether it could be observed in the context of a real-world current event – namely, the SARS-CoV2 (COVID-19) pandemic. We employed an individual differences design to investigate whether those who appraised COVID-19 as posing a greater threat to their health would also be more likely to morally condemn the actions of others for engaging in COVID-19 spreading behaviours. Representative samples – matched on age, gender and ethnicity to the most recent national census – were recruited in the U.K. and U.S. during the first year of the COVID-19 pandemic. In both samples we observed a significant positive relationship between COVID-19 threat-appraisals and moralization of others COVID-19 spreading behaviours (e.g. "never washing hands after being in public"). Importantly, these patterns of relationships remained robust even after controlling for participants' age, gender, self-reported social and economic conservatism and Big-Five Neuroticism (Soto & John, 2017).

7.1.2. **What are the potential functional benefits of moralization for the Individual**

The second core research question in this thesis was to examine how the psychological or behavioural outcomes of moralization were functional for the individual. To this end, in Chapter 6, we conducted a pre-post study around a highly polarized political event in the US. Specifically, we investigated individuals' degree of moralization and their perceptions of factual information developed around the release of Special Prosecutor Robert Mueller's Report on Investigations in the Russian Interference in the 2016 US election ("The Mueller Report"). We found that, people who had strongly moralized beliefs prior to the Mueller Report being released, were more likely to state – after the Report was released – that it had confirmed their expectations, and favoured their overall position. We also found that participants who perceived the report as favouring their overall position were more likely to report increases in Moralizaiton from Time 1 to Time 2.

7.2. **Theoretical Implications**

Across three empirical chapters, we found robust support for the relevance of individual threat-appraisal processes on the propensity to moralize others' actions (Chapters 3 to 5). In addition, we found that moralization was associated with an increased desire for cultural tightness (Chapter 5) and led to increased propensity among participants to view new information as supporting their moralized beliefs (Chapter 6). This section will present a contextualisation of these findings within the current literature, and engage in a theoretical discussion of how moralization may be conceptualised as a functional response to threat-appraisal.

7.3.   Threat-Appraisal as a Motivator of Moralization

The clearest finding throughout the empirical chapters of this thesis was the existence of a relationship between threat-appraisal and moralization in the form of moral amplification of third-party actions. These findings suggest that individuals who feel threatened have a greater propensity to draw on moral significance when evaluating the actions of others.

Within the field of moral psychology, our findings build on a number of separate lines of research demonstrating that threat-appraisal of a specific context can lead to increased moralization (Jordan & Monin, 2008; Laham & Corless, 2016; Murray et al., 2019). For example, moralization in past work has been associated with threats to self-concept (Jordan & Monin, 2008), threat of disease (Henderson & Schnall, 2021; Murray et al., 2019; van Leeuwen et al., 2012), and threat of social exclusion (Pfundmair & Wetherell, 2019). It is difficult to definitively conclude that these separate lines of research represent the same mechanism of threat-appraisal to moralization that was observed in Chapters 3 to 5. However, our findings do point to the possibility that the functional element in each of these studies pertaining to moralization may have been the appraisal of threat, rather than the specific context of the threat.

Outside of the specific context of morality, our findings are also in line with a growing body of evidence in political psychology showing that experiences of threat across a wide range of contexts are associated with outcomes that are analogous to moralization. For example, analogous to findings that moralization leads to perceptions of beliefs as universal and beyond compromise (Skitka, 2010), threat-appraisals across a wide range of domains has also been found to predict willingness to adopt extreme ideological positions (Haas & Cunningham, 2014; Jonas et al., 2014; Nash et al., 2011), cultural norms (van den Bos et al., 2005) and reductions in support for political compromise (Haas, 2016). Seen in this context, our findings that threat-

appraisals are reliably associated with moralization point to the possibility that moralization may be a relevant intermediary process between individuals' threat-appraisal and their adoption of extreme belief systems.

7.3.1. **Threat-Appraisal/Moralization Link not Explained by Disgust**

As the emotion that has most often been associated with morality and moralization (Landy & Goodwin, 2015), disgust has played a prominent role in both empirical research and theorising around moralization (Kupfer & Giner-Sorolla, 2017; Rozin et al., 1999; Schnall et al., 2008). It was therefore important to establish that any observed statistical relationship between threat-appraisal constructs and moralization could not simply be accounted for by their shared variance with disgust.

In Study 3 of Chapter 3, we found that the relationship between threat-sensitivity and moralization was partially accounted for by participants' experience of disgust at reading the vignettes. However, we also found that participants' experience of threat-appraisal emotions (i.e. anxious, worried, troubled) remained a significant mediator of the relationship between threat-sensitivity and moralization after disgust was controlled for. In Chapter 4, we found that participants' experience of threat-appraisal emotions in everyday life was a significant predictor of moralization of third-party behaviours, while participants' experience of disgust was not a significant predictor.

While these findings do not rule out the relevance of disgust in the moralization process. They provide strong support for the inference that feelings of disgust alone cannot account for the relationship between threat-appraisal and moralization that has been consistently observed in this thesis. This finding corroborates a growing body of evidence questioning the centrality of disgust as a unique moral emotion (Landy & Piazza, 2019; Wisneski et al., 2020).

7.3.2. **Social Conservatism does not explain the Threat/Moralization link**

Importantly, both Chapters 3 and 5 found that the relationship between threat-appraisal and moralization could not be accounted for by participants' self-reported levels of social conservatism. Indeed, in Chapter 5, we found a significant negative relationship between social conservatism and both threat-appraisal of COVID-19 and moralization of COVID-19 spreading behaviours in the US; we also found that the relationships between these same variables were non-significant in the UK. Such findings point to the importance of social and political context in determining the relevance of ideology in the dynamics of threat-appraisal and moralization. Our findings thus challenge a number of past accounts which suggest that threat is particularly relevant to the adoption of conservative ideology (Janoff-Bulman & Carnes, 2013; Thórisdóttir & Jost, 2011). Our more general threat-based account thus brings theorizing in the moral domain in line with more recent evidence challenging the notion that conservatives are dispositionally more sensitive to threat (Bakker et al., 2020; Crawford, 2017).

7.3.3. **Incidental Effect of Threat Appraisal?**

In Chapter 4 we observed that participants' real-world emotional experiences previously related to threat-appraisal (i.e. worried, anxious, troubled) were significant predictors of moral amplification of participants' judgments. Importantly, we found that participant's experience of these emotions prior to being presented with the vignettes, nonetheless influenced their judgments of the moral wrongness of the actions portrayed therein.

These findings speak to current uncertainties in the moral psychology literature regarding the relationship between incidental emotions and moral judgment (Landy & Goodwin, 2015; Wisneski et al., 2020). A number of early influential studies reported robust evidence that the induction of feelings of disgust (e.g. by fart spray) led to increases in the general harshness of

moral judgments for unrelated actions (see Chapter 1). However, the replicability of these findings have recently come into question, with both large-scale meta-analyses and multi-lab replication studies failing to find evidence to support the existence of a relationship between incidental disgust and moral amplification (Ghelfi et al., 2020; Jylkkä et al., 2021; Landy & Goodwin, 2015). These failures to replicate the highly influential line of research regarding incidental disgust have led to broader theoretical uncertainties regarding the validity of the influential Social Intuitionist Theory (Haidt, 2001) of morality.

According to the Social Intuitionist Theory (SIT: Haidt, 2001) moral judgments are most often the product of unconscious and automatic processes, as opposed to rational thinking. Thus, rather than consciously deliberating on relevant moral principles prior to making a moral judgment, individuals will often rely on heuristic cues of morality that they have acquired through social reinforcement. These cues often take the form of certain emotions (e.g. disgust) that are often experienced in the context of morality, thus allowing the individual to intuit that experiencing disgust while judging another's actions signals that the action is morally wrong.

The findings of Chapter 4 appear to partially support the SIT perspective. Specifically, our findings suggest that individuals may indeed make moral judgments based on unconscious or automatic processes. However, contrary to SIT, it appears unlikely that the *incidental* effect of threat-appraisal emotions on moral judgment observed in Chapter 4 were driven by participants using the experience of these emotions to intuit moral wrongness. This is particularly so, given that the emotions themselves (anxious, worried, troubled) are not traditionally considered morally relevant (Landy & Piazza, 2019).

An alternative explanation for our findings may be that the experience of threat-appraisal emotions prime individuals to moralize. In other words, it is possible that the incidental effects of

certain emotions on moralization are not because they signal morality, but rather because they place the individual in a state of moral vigilance (Murray et al., 2019). A growing body of work suggests that individuals rely on judgments of moral character in others as heuristic cues for whether they can be trusted, or pose a threat (Brambilla et al., 2018, 2021; Klebl et al., 2021). Applying these insights to the current context, it is possible that the experience of threat-appraisal emotions may prime individuals to look for cues of moral character, as a means of detecting individuals in their environment who are particularly likely to pose a threat.

7.4. Decreased Uncertainty as an Outcome of Moralization

In Chapter 6, we found that moralization was not only associated with greater confidence about future events, but also influenced evaluation of new information. Specifically, we found that individuals with highly moralized beliefs were more likely to perceive new information as confirming their prior expectations, and favouring their overall position on the moralized topic. These findings suggest that, when beliefs are highly moralized, they may be used by individuals as a means of interpreting new factual information. These findings are broadly consistent with work showing that people often attribute objective factual status to their moralized beliefs. In other words, individuals often believe it is possible to hold an objectively right or wrong position on topics that they have moralized (Goodwin & Darley, 2008, 2012; Stanley et al., 2020). However, as far as we are aware, the findings in Chapter 6 are one of the first to demonstrate that moralized beliefs can also influence the evaluation of factual information.

7.5. Evaluating the Functional Action-Based Moralization Account

The action-based moralization account was developed as a functional framework for systematically identifying when and why individuals may be motivated to engage in moralization. The core premise of the action-based moralization account is that individuals are

motivated to moralize as a means of facilitating certainty about future action. More specifically, the account starts with the premise that threat-appraisals induce a state of anxious-uncertainty that prevents future goal oriented behaviour until the threat is neutralised, which in turn leads to a heightened motivation to quickly resolve uncertainty. Critically, the action-based account proposes that this heightened motivation to resolve uncertainty increases the likelihood of moralization, a process which activates a series of behavioural and cognitive schemas that facilitate the reduction of uncertainty and thus facilitate future action.

The action-based account of moralization puts forward three mechanisms by which the process of moralization allows individuals to overcome uncertainty and engage in more efficient action: (1) moralization facilitates decision-making by creating a dramatic spreading of alternatives between options that are relevant to the moralized belief; (2) moralization allows individuals to maintain a sense of lowered uncertainty by creating an internal epistemic authority that is resistant to disconfirmation by new information; and (3) moralization facilitates the development of a shared-reality that allows the individual to draw on other ingroup members to respond to new threats.

The empirical chapters of this thesis sought to evaluate the feasibility of the action-based moralization account, by testing two of its core premises: (1) that moralization is likely to occur in response to threat-appraisal which induces a state of anxious-uncertainty; and (2) moralization reduces uncertainty. Across four empirical chapters, we found broad support for these two premises

Specifically, we found that individuals who have a propensity to appraise their environment as threatening were more likely to moralize third-party actions, and further that this threat-sensitivity/moralization relationship was mediated by the experience of threat-appraisal

emotions (anxious, worried, troubled). We also found that real-world changes in the experience of these threat-appraisal emotions predicted participants' tendency to moralize third-party behaviours. Finally, we found across two representative samples, that variance in threat-appraisal of the COVID-19 pandemic were related to moralization of COVID-19 spreading third-party behaviours. These findings provide initial phenomenological evidence of the relevance of threat-appraisal to moralization, and thus support the basic premise of the action-based moralization account that moralization responds specifically to a state of threat-induced anxious-uncertainty that individuals are highly motivated to resolve.

In the final empirical study of this thesis, we also found that participants' moralization of their position on a political issue increased their confidence about future events, and also influenced their evaluation of new information in a way that allowed them to maintain their current beliefs. These findings provide initial support for the premise that moralization enables sustained reduction in uncertainty.

Taken together, the empirical studies presented in this thesis provide initial support for the feasibility of an action-based moralization account, and thus suggests that further research of moralization through this framework is warranted.

7.6.   Limitations and Future Directions

This thesis was a preliminary examination of the feasibility of a novel account of moralization. It is important to emphasise that the objective of the thesis was not to substantiate every premise of the action-based moralization account, but rather to test its merit as a potential avenue of inquiry into the functions and motivations of moralization. For this reason, the empirical work presented in this thesis contains a number of limitations in terms of their ability to support inferences about the action-based moralization account.

Perhaps the most notable limitation of the empirical evidence presented in this thesis is that it did not seek to directly evaluate the relevance of uncertainty in the relationship between threat-appraisal and moralization.

Second, while a significant strength of the thesis was that the methods applied in the empirical chapters were particularly conducive of ecological validity and generalizability, these methods were not capable of providing strong evidence of causal direction. Specifically, while the thesis provides strong evidence of the general relationship between threat-appraisal and moralization across multiple contexts, it provides less direct evidence that threat-appraisal causes moralization.

Third, because the aim of this thesis was to evaluate the initial feasibility of the action-based moralization account, a number of its core premises remain untested. Notably, there is little direct evidence in this thesis on whether threat-induced moralization leads to a lowered state of uncertainty.

7.7. Conclusion

The past 20 years of investigation into subjective morality has revealed that moral change is a complex and multifaceted phenomenon that is capable of producing cognitions and behaviours with both constructive and destructive social and political consequences. The current work builds on this insight by proposing and testing a theory of moralization that frames the process in terms of how it interacts with factors in the individual's physical and social context. It is my hope that this approach to moralization theory will serve as the a conceptual bridge by which insights from multiple disciplines can be brought in to better understand the process of moral change.

References

Brambilla, M., Biella, M., & Freeman, J. B. (2018). The influence of visual context on the evaluation of facial trustworthiness. *Journal of Experimental Social Psychology*, *78*, 34–42. https://doi.org/10.1016/j.jesp.2018.04.011

Brambilla, M., Sacchi, S., Rusconi, P., & Goodwin, G. P. (2021). Chapter Four - The primacy of morality in impression development: Theory, research, and future directions. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 64, pp. 187–262). Academic Press. https://doi.org/10.1016/bs.aesp.2021.03.001

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 U.S. Presidential Election Campaign. *Journal of Social and Political Psychology*, *3*(2), 211–237. https://doi.org/10.5964/jspp.v3i2.434

Caseras, F. X., Fullana, M. A., Riba, J., Barbanoj, M. J., Aluja, A., & Torrubia, R. (2006). Influence of individual differences in the Behavioral Inhibition System and stimulus content (fear versus blood-disgust) on affective startle reflex modulation. *Biological Psychology*, *72*(3), 251–256.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1178–1198. https://doi.org/10.3758/s13428-014-0551-2

Cools, R., Calder, A. J., Lawrence, A. D., Clark, L., Bullmore, E., & Robbins, T. W. (2005). Individual differences in threat sensitivity predict serotonergic modulation of amygdala response to fearful faces. *Psychopharmacology*, *180*(4), 670–679.

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of

moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology*, *117*(1), 50–72. https://doi.org/10.1037/pspa0000149

Ghelfi, E., Christopherson, C. D., Urry, H. L., Lenne, R. L., Legate, N., Ann Fischer, M.,

Wagemans, F. M. A., Wiggins, B., Barrett, T., Bornstein, M., de Haan, B., Guberman, J., Issa, N., Kim, J., Na, E., O'Brien, J., Paulk, A., Peck, T., Sashihara, M., … Sullivan, D. (2020). Reexamining the Effect of Gustatory Disgust on Moral Judgment: A Multilab Direct Replication of Eskine, Kacinik, and Prinz (2011). *Advances in Methods and Practices in Psychological Science*, *3*(1), 3–23. https://doi.org/10.1177/2515245919881152

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism.

Cognition, *106*(3), 1339–1366. https://doi.org/10.1016/j.cognition.2007.06.007

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more

objective than others? *Journal of Experimental Social Psychology*, *48*(1), 250–256. https://doi.org/10.1016/j.jesp.2011.08.006

Haas, I. J. (2016). The Impact of Uncertainty, Threat, and Political Identity on Support for

Political Compromise. *Basic and Applied Social Psychology*, *38*(3), 137–152. https://doi.org/10.1080/01973533.2016.1169181

Haas, I. J., & Cunningham, W. A. (2014). The Uncertainty Paradox: Perceived Threat Moderates

the Effect of Uncertainty on Political Tolerance. *Political Psychology*, *35*(2), 291–302.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral

judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Henderson, R. K., & Schnall, S. (2021). Disease and Disapproval: COVID-19 Concern is

    Related to Greater Moral Condemnation. *Evolutionary Psychology*, *19*(2),

    14747049211021524. https://doi.org/10.1177/14747049211021524

Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., Nash, K., Proulx, T.,

    & Quirin, M. (2014). Threat and Defense. In *Advances in Experimental Social*

    Psychology (Vol. 49, pp. 219–286). Elsevier. https://doi.org/10.1016/B978-0-12-800052-

    6.00004-4

Jordan, A. H., & Monin, B. (2008). From Sucker to Saint:Moralization in Response to Self-

    Threat. *Psychological Science*, *19*(8), 809–815. https://doi.org/10.1111/j.1467-

    9280.2008.02161.x

Jylkkä, J., Härkönen, J., & Hyönä, J. (2021). Incidental disgust does not cause moral

    condemnation of neutral actions. *Cognition and Emotion*, *35*(1), 96–109.

    https://doi.org/10.1080/02699931.2020.1810639

Klebl, C., Rhee, J. J., Greenaway, K. H., Luo, Y., & Bastian, B. (2021). Physical Attractiveness

    Biases Judgments Pertaining to the Moral Domain of Purity. *Personality and Social*

    Psychology Bulletin, 01461672211064452. https://doi.org/10.1177/01461672211064452

Koval, P., Hinton, J., Dozo, N., Gleeson, J., Alvarez, M., Harrison, A., Vu, D., Susanto, R.,

    Jayaputera, G., & Sinnott, R. (2019). SEMA3: Smartphone Ecological Momentary

    Assessment, Version 3. *Computer Software]. Retrieved from Http://Www. Sema3. Com*.

Kruepke, M., Molloy, E. K., Bresin, K., Barbey, A. K., & Verona, E. (2018). A brief assessment

    tool for investigating facets of moral judgment from realistic vignettes. *Behavior*

    Research Methods, *50*(3), 922–936. https://doi.org/10.3758/s13428-017-0917-3

Kupfer, T. R., & Giner-Sorolla, R. (2017). Communicating moral motives: The social signaling function of disgust. *Social Psychological and Personality Science*, *8*(6), 632–640.

Laham, S. M., & Corless, C. (2016). *Threat, morality and politics: A differentiated threat account of moral and political values*.

Landy, J. F., & Goodwin, G. P. (2015). Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence. *Perspectives on Psychological Science*, *10*(4), 518–536. https://doi.org/10.1177/1745691615583128

Landy, J. F., & Piazza, J. (2019). Reevaluating Moral Disgust: Sensitivity to Many Affective States Predicts Extremity in Many Evaluative Judgments. *Social Psychological and Personality Science*, *10*(2), 211–219. https://doi.org/10.1177/1948550617736110

Lissek, S., & Powers, A. S. (2003). Sensation seeking and startle modulation by physically threatening images. *Biological Psychology*, *63*(2), 179–197.

Murray, D. R., Kerry, N., & Gervais, W. M. (2019). On Disease and Deontology: Multiple Tests of the Influence of Disease Threat on Moral Vigilance. *Social Psychological and Personality Science*, *10*(1), 44–52. https://doi.org/10.1177/1948550617733518

Nash, K., McGregor, I., & Prentice, M. (2011). Threat and defense as goal regulation: From implicit goal conflict to anxious uncertainty, reactive approach motivation, and ideological extremism. *Journal of Personality and Social Psychology*, *101*(6), 1291–1301. https://doi.org/10.1037/a0025944

Pfundmair, M., & Wetherell, G. (2019). Ostracism drives group moralization and extreme group behavior. *The Journal of Social Psychology*, *159*(5), 518–530. https://doi.org/10.1080/00224545.2018.1512947

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping

    between three moral emotions (contempt, anger, disgust) and three moral codes

    (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*(4),

    574–586. https://doi.org/10.1037//0022-3514.76.4.574

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as Embodied Moral

    Judgment. *Personality and Social Psychology Bulletin*, *34*(8), 1096–1109.

    https://doi.org/10.1177/0146167208317771

Stanley, M. L., Marsh, E. J., & Kay, A. C. (2020). Structure-seeking as a psychological

    antecedent of beliefs about morality. *Journal of Experimental Psychology: General*,

    149(10), 1908–1918. https://doi.org/10.1037/xge0000752

van den Bos, K., Poortvliet, P. M., Maas, M., Miedema, J., & van den Ham, E.-J. (2005). An

    enquiry concerning the principles of cultural norms and values: The impact of uncertainty

    and mortality salience on reactions to violations and bolstering of cultural worldviews.

    Journal of Experimental Social Psychology, *41*(2), 91–113.

    https://doi.org/10.1016/j.jesp.2004.06.001

van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen

    prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human*

    Behavior, *33*(5), 429–437. https://doi.org/10.1016/j.evolhumbehav.2011.12.005

Wisneski, D. C., Hanson, B. E., & Morgan, G. S. (2020). The roles of disgust and harm

    perception in political attitude moralization. *Politics & the Life Sciences*, *39*(2), 215–227.

    https://doi.org/10.1017/pls.2020.22

Zheng, Y., Xu, J., Jia, H., Tan, F., Chang, Y., Zhou, L., Shen, H., & Qu, B. (2011).

    Electrophysiological correlates of emotional processing in sensation seeking. *Biological*

    Psychology, *88*(1), 41–50.

    Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation*

    Seeking. Cambridge University Press.

De Weerd, M., & Klandermans, B. (1999). Group identification and political protest: farmers'
protest in the Netherlands. *European Journal of Social Psychology, 29*(8), 1073-1095.
https://doi.org/10.1002/(SICI)1099-0992(199912)29:8<1073::AID-EJSP986>3.0.CO
2-K

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of
moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.

Haidt, J., & Graham, J. (2007, 2007//). When Morality Opposes Justice: Conservatives Have
Moral Intuitions that Liberals may not Recognize. *Social Justice Research, 20*(1), 98-
116. https://doi.org/10.1007/s11211-007-0034-z

Klandermans, B. (1984). Mobilization and Participation: Social-Psychological Expansisons of
Resource Mobilization Theory. *American Sociological Review, 49*(5), 583-600.

Ryan, T. J. (2017). No compromise: Political consequences of moralized attitudes. *American
Journal of Political Science, 61*(2), 409-423.

Skitka, L. J. (2010). The Psychology of Moral Conviction. *Social and Personality Psychology
Compass, 4*(4), 267-281. https://doi.org/10.1111/j.1751-9004.2010.00254.x

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral Conviction: Another Contributor to
Attitude Strength or Something More? *Journal of Personality and Social Psychology,
88*(6), 895-917. https://doi.org/10.1037/0022-3514.88.6.895

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting
inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and
Cognition*.

Turner-Zwinkels, F. M., Zomeren, M., & Postmes, T. (2017). The moral dimension of politicized
identity: Exploring identity content during the 2012 Presidential Elections in the USA.
British Journal of Social Psychology, 56(2), 416-436.