

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Psychology

2025

The Cognition of Social Norms: A Process Theory Investigation of Social Norm Perceptions and Interventions

Composto, Jordana

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp016q182p56q>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

THE COGNITION OF SOCIAL NORMS: A PROCESS THEORY INVESTIGATION OF
SOCIAL NORM PERCEPTIONS AND INTERVENTIONS

Jordana Rose Winey Composto

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY

THE DEPARTMENT OF PSYCHOLOGY

Adviser: Elke U. Weber

May 2025

© Copyright by Jordana Rose Winey Composto, 2025. All rights reserved.

Abstract

Accurate understanding of the social world is vital for addressing complex societal challenges. As humans navigate information-rich environments, they form perceptions about social norms that critically influence collective action on existential threats like climate change and democratic erosion. These perceptions often contain systematic distortions: underestimating public support for climate policies and overestimating endorsement of antidemocratic actions. While previous research has documented these misperceptions and their consequences, my work investigates their underlying cognitive mechanisms. In the first chapter, I present a meta-analysis of query theory, finding support for its central claims that stimulus framing affects the order in which people evaluate options which, in turn, affects the balance of support and leads to choice anomalies. In the second chapter, I extend query theory to social norm perception, developing a memory-retrieval theory. Norm perceptions are, in part, constructed by the order and content of the memory-querying process. The third chapter investigates how information environments, particularly media coverage, shape perceptions of democratic norms in the United States. I use large-language models to identify and quantify the amount of social norm information communicated through cable news. Exposure to norm information in one's media diet is associated with attitudes and partisan norm perceptions about democracy. The fourth chapter examines the cognitive mechanisms through which social norm information influences policy support across energy technology domains. Query order and content are affected by social norm information, impacting policy support. This finding holds for participants who are not skeptical of the policy or norm information. This multi-pronged investigation advances our understanding of how social perceptions are constructed and affect complex societal transitions, addressing critical challenges that require coordinated collective action based on accurate perceptions of the normative landscape.

Table of Contents

Abstract	3
Table of Contents	4
Acknowledgements	5
Introduction	7
Chapter 1: A Meta-Analysis of Query Theory, a Psychological Process Account of Framing Effects	14
Chapter 2: A memory-retrieval perspective on the construction of social norm perceptions	39
Chapter 3: Social Norms in Media: How the Prominence of Social Norms in Cable Television Shapes Norm Perceptions and Attitudes about Democracy	87
Chapter 4: A process theory of social norm information interventions	123
General Discussion	168
References	172
Appendices	197

Acknowledgements

None of this would have been possible without the support, encouragement, and intellectual stewardship of my advisor, Elke Weber. Thank you for challenging me to follow my ambitions and to focus on impact-driven research without compromising theoretical rigor. Your interdisciplinary approach has been instrumental in inspiring me to forge meaningful connections between scientific inquiry, practical applications, and policy implications. I hope to never stop learning from you.

Thank you to Alin Coman, my secondary advisor, for teaching me how to think deeply and showing me the power of slowing down. Thank you for taking the time to think through ideas with me and for giving me an academic home in the psychology department.

Thank you to the many people who have played an indispensable role in my academic career. To the faculty who helped me navigate graduate school, follow my side quests, and develop impact-driven research questions—Chris Greig, Casey Lew-Williams, and Nicole Shelton—thank you for showing me what is possible. Thank you to Alin Coman, Betsy Levy Paluck, Anuj Shah, Nicole Shelton, and Natalia Vélez, for serving on my committees and for your invaluable feedback throughout the dissertation process. To Sara Constantino, thank you for chasing down ideas with me and for taking me seriously so early on. To Jordan Starck, thank you for welcoming me onto your project and for modeling impactful, interdisciplinary research. To Russ Walters and Jodi Mindell, thank you for introducing me to research.

Thank you to the collaborators, lab mates, and research assistants who made this dissertation not only possible but enjoyable. To Kris Nichols, the best type of collaborator is one who is a little overambitious and crazy enough to make it work. To Roh Majumdar, thank you for going first. To Johanna Matt-Navarro, thank you for being a sanity check on all things. To my

collaborators on this dissertation, Eric Johnson, Aya Salim, Eva Lu, Helena Grosse, and Shannon Duncan, thank you for your work and your time. To all my lab mates in the Behavioral Science for Policy Lab, Cognition in Collectives Lab, and the Center for Research on Environmental Decisions in Berlin, thank you for inspiring and laughter-filled lab meetings each week. Thank you to the friends that graduate school brought into my life—Stats Attwood, Isaac Breinyn, Kara Enz, Sana Khan, Leon Mait, Asana Okocha, Yeji Park, Logan Pearce, Joanna Schneider, Evan Underhill, and Renée Waters.

Thank you to the friends who make my life exciting and wonderful. To Bri Williams, Katie Papa, Ayana Bowman, Izzi Levy, and the Philadelphia Eagles, you mean the world to me.

To my family, thank you for your fierce encouragement and support. To Rebecca Composto, the greatest sister in the world, for knowing when to call. To my dad, Russ Composto, for teaching me to take one step forward every day and modeling what it means to care for students. To my mom, Karen Winey, for raising me as a list-maker and teaching me that “no” is a complete sentence. To Mimi and Pop-pop, for cheering on every part of my life. And to Chris Browne, thank you for choosing to build a life with me and for reminding me to drink water. I love being the social scientist in a family of engineers.

Introduction

Understanding the social world is fundamental to living in a society. Humans are born into cultures with complex information systems that exceed individual capacity (Henrich, 2016). As social beings, humans have evolved to exceed that individual capacity by processing complex informational signals from individuals, groups, and institutions to understand what is expected or permissible in different contexts. These evolutionary adaptations that have equipped us with social emotions and social cognition help us determine whom to trust, assess environmental safety, and interpret subtle social cues in our daily interactions. In recent decades, the growth and proliferations of social media platforms and traditional media channels have increased the volume of social cues that inform social norms. Social norms represent attitudes or behaviors that are either considered desirable (injunctive) or prevalent (descriptive) within specific groups or contexts (Cialdini & Goldstein 2004; Legros & Cislighi 2020; Hector & Opp, 2001). Norms extend beyond interpersonal dynamics to shape our perceptions of institutions, influencing our confidence in democratic systems, expectations of corporate behavior, and attitudes toward governance structures (Otto et al., 2020). Social norms become particularly crucial when addressing complex challenges like climate change, where significant temporal and spatial separation exists between actions and their subsequent outcomes.

Accurate perception of social norms enables individuals to participate effectively in civic life, resist self-silencing, and contribute to solving collective problems. When these perceptions become distorted, we risk entrenched polarization, institutional mistrust, and coordination failures at moments when cooperation is most vital. Our ability to address existential threats thus depends significantly on perceiving and responding to the true normative landscape rather than distorted or incomplete versions that impede progress.

The relationship between social norms and policy is reciprocal and dynamic. Social norms about policy sit at the intersection of individual and collective solutions to societal challenges, with signals from the public shaping governance and technological strategy in the energy transition (Otto et al., 2020). Citizens communicate policy preferences both directly through public discourse—signing petitions or contacting representatives—and indirectly by voting for aligned candidates. Simultaneously, institutions and elites shape norm perceptions through media discourse that adjusts expectations of others' beliefs, attitudes, or behaviors (Nyborg et al 2016; Clark et al., 2023). This dissertation focuses specifically on perceived norms affecting complex societal threats: energy transition policy and threats to democracy.

Both climate change and attitudes towards democracy are characterized by profound uncertainty and widespread public misconceptions regarding the actual social norms. They are extensively deliberated in both public discourse and private conversations—the potential sources of normative information are numerous and frequently contradictory. Empirical evidence indicates that individuals systematically underestimate public support for climate policies (Sparkman et al., 2022) and overestimate public endorsement of antidemocratic actions (Kim et al., 2022). These misperceptions can precipitate self-censorship and retaliatory behavior, both of which contribute to system destabilization.

When public knowledge regarding a policy domain, such as novel energy technologies, is limited, individuals rely more on the perceived attitudes of others. Underestimated support for climate policy attenuates expressions of genuine support. Conversely, when the public overestimates support for antidemocratic actions, trust in individuals and institutions may be compromised due to exaggerated perceptions regarding the erosion of democratic norms. While the extant literature has predominantly focused on the causal factors underlying these

misperceptions, this dissertation investigates the cognitive mechanisms through which these phenomena manifest by applying a process framework to the perception and formation of norms that affect these complex social systems challenges.

Chapter 1 presents a meta-analytic evaluation of query theory, a psychological process theory that explains how attentional processes and memory dynamics influence preference construction and decision-making (Johnson et al., 2007; Weber et al. 2007). The chapter provides the first comprehensive quantitative assessment of query theory since its inception. I use meta-analytic methods to evaluate the core premises of query theory. The first analysis examines how decision frames affect query order, analyzing studies that employ a query theory paradigm to elicit thought listings while investigating the relationship between decision framing and judgments or choices. The second analysis evaluates the mediating mechanisms of query theory by analyzing the proportion of the total effect that is mediated by query content and query order. The third analysis tests the causal claim that query order affects choice by comparing the effects of natural versus unnatural query order on decision outcomes.

The findings from these meta-analyses, synthesizing data from 27 papers with 14,602 total participants, provide substantial support for query theory's central claims. Specifically, the results demonstrate that decision frames significantly affect query order with a medium effect size ($d = 0.34$), that query mechanisms partially mediate the relationship between decision frame and choice, and that manipulating query order substantially reduces the effect of decision framing on choice, decreasing the effect size from $d = 0.92$ to $d = 0.39$. These results validate the query theory process theory of framing effects on judgment and decision anomalies that are typically modeled by Prospect Theory's functional account (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

Chapter 2 extends query theory from individual preference construction to the domain of social cognition, specifically focusing on the constructed perception of social norms. The chapter is motivated by growing evidence of pluralistic ignorance regarding environmental policies and behaviors, where individuals systematically underestimate public support for pro-environmental initiatives (Geiger & Swim, 2016; Sparkman et al., 2022). This underestimation creates false social realities that lead to self-silencing and impede collective action on critical issues such as climate change. I argue that while previous research has documented these perception gaps and their consequences, less attention has been paid to the underlying cognitive processes that construct inaccurate social norm perceptions. Understanding these cognitive mechanisms is crucial as it explains both how inaccurate perceptions initially form and why amplify or become entrenched. A better understanding of the cognition of social norm perceptions can provide the foundation for developing more effective interventions that can correct misperceptions and prompt accurate social understanding.

The chapter develops a memory-retrieval theory of norm perception by integrating insights from Kahneman and Miller's norm theory (1986) with query theory. Through two studies, I investigate whether social norm perceptions are constructed through memory-retrieval processes, identify which features of memory predict norm perceptions, and evaluate norm perceptions about corporate responsibility and individual food choices. Study 1 serves as an exploratory analysis examining how the stimulus frame (asking for norm agreement or disagreement) and norm type (injunctive versus descriptive) affect query order and content. Study 2 formally tests four hypotheses: that question frame affects query order and content, that norm type influences query order and content, that query order and content predict norm perception, and that norm type affects

the types of memories (semantic versus episodic) that individuals access when forming norm perceptions.

This research represents an innovative theoretical extension of query theory to social cognition, suggesting that social norm perceptions are constructed at least partially through memory-retrieval processes and are influenced by the order and content of retrieved memories. By identifying the cognitive mechanisms underlying norm perception, this chapter contributes to our understanding of how individuals develop perceptions about social consensus, particularly regarding environmental issues.

Chapter 3 building on the ideas of norm construction by examining how the information environment, particularly media coverage, shapes perceptions of democratic norms in the United States. There is a general consensus that American democracy is under threat from political sectarianism (Finkel et al., 2020) and the weakening of democratic norms (Holliday et al. 2024). While empirical evidence suggests that antidemocratic attitudes remain primarily confined to political elites rather than the general public (Holliday et al. 2024), high levels of affective polarization contribute to misperceptions about public endorsement of antidemocratic norms (Iyengar et al., 2019). I investigate the social norm information communicated through cable television news and the effect on individual attitudes and social norm perceptions about democracy.

The chapter explores the real, information-dense environment through which people receive critical information, both knowingly and unknowingly, about democracy. Media polarization, particularly in prime-time television, has been linked to increased viewer partisanship and engagement in political discord. While individuals attempt to identify social patterns from this information deluge, their perceptions are often shaped by the ideological slant of their news

sources, which affects their understanding of what behaviors are permissible within a democratic system.

I leverage recent advances in large language models (LLMs) to investigate social norm information in cable television news and how media consumption affects attitudes and social norm perceptions about antidemocratic actions. This chapter has three sections: first, I evaluate how well different LLMs can identify norm information in cable news transcripts; second, I present descriptive analyses of cable news coverage in the two months preceding the 2024 United States Presidential Election; and third, I conduct a survey of registered voters to understand how norm prominence within their media diet shapes their attitudes and perceptions about democracy. By linking participants' reported media consumption to my text corpus, I quantify exposure to democracy coverage and investigate whether norm prominence in media diet correlates with democratic attitudes and perceptions.

Chapter 4 investigates the cognitive mechanisms through which social norm information influences policy support. Building on established research demonstrating that social norms shape behaviors, attitudes, and beliefs, this chapter focuses on the question of how norm information influences policy support at a cognitive level. I employ a naturalistic type-aloud procedure developed in query theory to identify the cognitive mechanisms that explain the effects of norm information interventions.

The chapter presents a series of studies that test a query theory explanation of social norm interventions in various policy contexts. Study 1 examines both positive and negative norm provision in a hypothetical policy scenario. Studies 2a-c test hypotheses in real policy contexts with accurate norm information, each addressing a different policy domain: energy efficiency with positive norm treatment, labor organizing with positive norm treatment, and energy technology

with negative norm treatment. Study 3 manipulates query order to test the causality of the query theory explanation using the same policy and norm context as Study 2c. Finally, Study 4 extends the findings to a more naturalistic information environment by providing policy context and norm information through actual television news clips related to support for large-scale solar farms.

I test four primary hypotheses across these studies: that social norm information affects policy support (with positive norms increasing support and negative norms decreasing it); that norm information influences query order and content; that query order and content mediate the effect of social norm information on policy support; and that manipulating query order to an "unnatural" sequence attenuates the effect of norm information on policy support and query content. By investigating these hypotheses, the chapter provides insights into the cognitive processes through which norm information shapes policy attitudes, with implications for designing more effective and scalable normative interventions.

This dissertation presents a multi-pronged investigation of how cognitive processes, particularly memory-retrieval mechanisms, influence individual decision-making, social norm perception, and policy attitudes. The research progresses from a meta-analytic validation of query theory in individual preference construction (Chapter 1) to its extension to social norm perception (Chapter 2), followed by examinations of how news media shapes democratic norm perceptions (Chapter 3) and how norm information affects policy support through cognitive mechanisms (Chapter 4). This work advances our understanding of the psychological processes underlying social cognition of diverse information sources in contexts of complex societal change.

Chapter 1: A Meta-Analysis of Query Theory, a Psychological Process Account of Framing Effects^{1 2}

Prospect theory (PT), introduced by Daniel Kahneman and Amos Tversky in 1979 and further refined in 1992 (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), is probably psychology's most successful export to economics and other social sciences. It was prominently cited as a 'formulation of a new branch of economics' in the citation awarding the Nobel Prize for Economic Science to Kahneman in 2002. Its insights modify expected utility theory (Von Neumann & Morgenstern, 1953), economics' normative model of risky choice. PT's most important insight is the observation that the outcomes of risky choice options are evaluated relative to a reference point rather than in an absolute fashion. Kahneman and Tversky modify expected utility theory by providing it with a utility function that is risk averse for perceived gains, risk seeking for perceived losses, and steeper for losses than for gains, a property described as loss aversion.

The addition to a flexible reference point, in conjunction with different functional mappings from objective outcomes to perceived value when the outcome is encoded as a relative loss or relative gain, allows PT to account for a broad range of phenomena collectively labeled framing effects that have been documented in both lab and real-world settings (Kahneman & Tversky, 2000). Framing effects are significant changes in the choice made as the result of a normatively irrelevant change in the description of choice options, e.g., by making one or another option the prechecked default option (Johnson & Goldstein, 2003) or by asking people to choose

¹ This chapter is currently under review at the *Journal of Risk and Uncertainty* for a Special Issue in Honor of Daniel Kahneman.

² This chapter was written with coauthors: Shannon M. Duncan (University of Pennsylvania), Eric J. Johnson (Columbia University), and Elke U. Weber (Princeton University).

between a smaller immediate or larger later reward by either asking them whether they would want to delay the immediate reward or whether they would want to accelerate the later reward (Loewenstein, 1988). PT models these changes in preference between two (normatively equivalent) frames by assuming that the frame changes the implicit reference point to which the outcomes of the choice options are compared.

PT provides a convincing account for framing effects, which are not predicted in any way by expected utility theory. But like expected utility theory, PT is a functional-level (“as-if”) theory, not a psychological process theory. That is, PT does not assume that decision makers in fact perform the comparisons of outcomes to a reference point nor that their choices are based on differential emotional responses to perceived gains versus perceived losses (Weber & Johnson, 2006). Instead, choices are made in some way “as-if” these operations occurred.

When researchers have examined loss aversion as a psychological process claim about emotions, namely that “the aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount” (Kahneman & Tversky, 1979, p. 279), they have often shown no differences between the magnitudes of pleasure and pain (Charpentier et al., 2016; Mellers, 2000). This raises the question whether other psychological process models can give rise to the choice pattern predicted by PT’s loss-averse value function. Query theory (QT) offers a psychological step-by-step process account of how values are constructed or assembled (Li et al., 2015) and how choice anomalies like framing effects (Weber et al., 2007) and other judgement anomalies emerge as a result (Johnson et al., 2007). Weber and Johnson (2006) suggest that preferences are subject to the dynamics of memory retrieval, and that decisions are affected by memory and attention processes. QT relies on four premises: (1) that decision-makers decompose a choice or valuation into component queries (Collins & Michalski,

1989), (2) that these queries are asked sequentially due to cognitive processing limitations, (3) that earlier queries shape subsequent queries through output interference (Dempster, 1995) or retrieval-induced forgetting (Anderson et al., 2000), resulting in a more successful recall of arguments for choosing the first option considered, and (4) that query order is affected by the frames of a decision, which focus attention on a different initial option.

Process tracing is a broad class of methodological approaches in the cognitive sciences that all involve collecting data at multiple time points during a complex task or decision. Such methods include verbal protocols where participants “think aloud” during a task and their speech is measured (Burchardi & Penczynski, 2014; Ericsson & Simon, 1980; Highhouse, 1994), mouse tracking protocols where participants’ mouse clicks are measured (Costa-Gomes & Crawford, 2006; Payne et al., 1993), and eye tracking protocols where participants’ eye movements, and the fixation on the visual items they attend to, are measured (Stewart et al., 2016; Wang et al., 2010). QT is closest to verbal “think aloud” protocols, within which there is an important distinction between concurrent verbal reports and retrospective reports. See Ericsson and Simon (1980), Russo and colleagues (1989) and Ericsson and Simon (1993) for a discussion of this distinction. QT uses a typed concurrent report.

The ecological validity of these protocols is critical to evaluate—does measuring a participant’s thought process change that process in some way? A meta-analysis of the Ericsson-Simon protocol finds that “think aloud” process tracing methods (i.e., when participants are asked to verbalize their thoughts during a task) do not lead to reactivity nor alter task performance (Fox et al., 2011). The review cautions against directing participants to explain their decision or process as that can lead to reactivity (i.e., distorting information by accessing it). For a relevant review of the tradeoffs surrounding different process tracing methods see Schulte-Mecklenbeck and

colleagues (2017). Taken together, the literature suggests that process tracing, when applied appropriately, measures a naturalistic decision process (Russo et al., 1989). QT leverages the validated instructions from verbal process tracing protocols (i.e., “share what goes through your mind”) while enabling researchers to collect larger sample sizes.

The typical QT investigation (e.g., Johnson et al., 2007, Weber et al., 2007) provides evidence about hypothesized variables and processes through process-tracing methodology that utilizes think-aloud protocols. The protocol is often a type-aloud paradigm as most studies are conducted online with demographically and socio-economically diverse samples. A first study traces the output of hypothesized queries (“why choose Option A?”, “why choose Option B?”), the order in which the two queries are generated and the relative evidence generated to answer each query. This query clustering was also validated with latent semantic analysis, a data-driven form of text analysis (Johnosn et al., 2007) and reaction time comparisons (Weber et al., 2007). Mediation analysis determines whether query order and query content, in combination or separately, predict respondents’ choices under each frame and, more importantly, whether differences in query order and query content between frames predict differences in choices.

A second study often probes the causal role of the change in query order between frames, by directly manipulating query order. The question about why to choose option A and why to choose option B, are explicitly prompted, and the order manipulated. This produces a ‘natural order’ which is consistent with how people do the task naturally. They first generate arguments for one option and only subsequently generate arguments for the other option. In contrast, reversing this natural order, the ‘unnatural order’, researchers show a smaller framing effect.

These meta-analyses capture different aspects of these studies. In the first meta-analysis, Meta-Analysis 1, I examine if the two frames produce different query orders, a necessary condition

for query theory to operate. In Meta-Analysis 2, I examine if these differences mediate the effects of framing upon choice. These two analyses are based then, typically, on the first experiment of the series. In Meta-Analysis 3, I focus on the second kind of study that manipulates the order of queries explicitly and see if the two orders produce different framing effects. The logic is that the ‘natural order’ should produce a large framing effect, but that by reversing the two queries, producing an ‘unnatural order’, the framing effect will be substantially diminished. The analyses are conducted as separate meta-analyses because the claims rely on different statistics (while the studies in Meta-Analyses 1 and 2 overlap, they evaluate different statistics from different analyses) and the causality claim is investigated in a different study (the nature of Meta-Analysis 3, that query order is manipulated, is mutually exclusive with Meta-Analyses 1 and 2, where there is not order manipulation).

The QT literature has focused on explaining anomalies in decisions and judgement, such as the endowment effect (Johnson et al., 2007) and inter-temporal discounting (Appelt et al., 2011; Hardisty et al., 2013; Weber et al., 2007). Various studies apply QT to real-world decisions, such as Knoll and colleagues (2015) who examine when people choose to claim social security, Majd and colleagues (2017) who examine food decisions made by restricted eaters, and Poon and colleagues (2023) who examine multi-alternative choice in the context of consumer electronics. QT has also been applied to default paradigms to understand the psychological mechanism and evaluate effectiveness (i.e., Dinner et al., 2011; Duncan & Weber, n.d.).

Chapter Overview

QT (Johnson et al., 2007; Weber et al., 2007) was introduced in the literature in 2007 and there has been substantial uptake from psychologists studying a range of anomalous decision-

making phenomena. The goal of this meta-analysis is to take stock of the last seventeen years of QT research and revisit the main claims of the theory with meta-analytic evidence. There has not been a comprehensive literature review, or meta-analysis, of QT since its conception and this is the first paper to evaluate the literature comprehensively and quantitatively. The first aim was to examine the effect of decision frame (e.g., being endowed versus not being endowed) on query order (measured as differences in what is made available from memory) (Meta-Analysis 1). I included all studies that use a QT study paradigm to elicit thought listings while examining the effect of decision frame on judgement or choice between two outcomes. The second aim was to evaluate the QT mechanism by meta-analyzing the mediation effect of query content and query order on the relationship between decision frame and choice. For all papers that conduct a mediation analysis, I extract the mediation proportion, or the proportion of the total effect that is mediated by query content and query order. This is Meta-Analysis 2. The third aim was to evaluate the causal claim that query order affects choice. I compare the effect of query order (natural versus unnatural) on choice (Meta-Analysis 3). Across these three meta-analyses, I offer a comprehensive evaluation of the core QT premises.

Meta-Analytic Methods

I followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for meta-analysis reporting (Moher et al., 2009). See Figure 1.1 for the PRISMA search flow chart. Data and scripts for this project are available on the Open Science Framework: https://osf.io/qzw9g/?view_only=95aeb447e25e43cca10d6df47f69680b.

Search Strategy

To identify candidate studies, I conducted a systematic literature search for original, experimental studies that use query theory to explain a behavioral phenomenon or an anomaly in rational choice. The search followed the PRISMA guidelines (Moher et al., 2009; Figure 1.1). I conducted the literature search using a systematic database search (i.e., Web of Science, APA PsychInfo, Scopus, EBSCO Academic Search Ultimate), a grey literature search (i.e., Google Scholar, APA PsychEXTRA), outreach to relevant academic societies (i.e., Society of Judgement and Decision Making), and targeted outreach to relevant authors. The full search string was: [(“query theory” OR “query-theory”) AND Publication date: 2007-2024] OR [cites Johnson et al. 2007] OR [cites Weber et al. 2007]). To mitigate publication bias, no journal or publication filters were used and unpublished work was included (e.g., dissertations, conference papers). When there were multiple reports of the same study (i.e., a poster and a published paper), the published peer-reviewed version was used. Papers in all languages were considered, there was a single relevant paper that was not written in English and DeepL Translate was used for translation.

Inclusion Criteria

I have three sets of inclusion criteria for the three studies. Meta-Analysis 1 inclusion criteria were: (a) studies must appropriately apply the thought listing method before a choice or judgment (as described by Johnson et al., 2007 or Weber et al., 2007), (b) studies do not manipulate the order or contents of thoughts or aspects, (c) the decision task must be an individual-level choice or judgment, (d) the results must report the standardized median rank difference (i.e., the indicator for query order) or the raw data for this calculation must be provided, (e) studies must report the size of the effect of condition on choice, and (f) studies must be original and experimental.

Meta-Analysis 2 inclusion criteria were: (a) studies must appropriately apply the thought listing methodology before a choice or judgment, (b) studies do not manipulate the order or

contents of thoughts or aspects, (c) the decision task must be an individual-level choice or judgment, (d) the results must report a mediation analysis of how the inclusion of query order and/or query content mediated the relationship between framing condition and choice or there must be raw data available to conduct this analysis, (e) when such a mediation analysis, studies must report a direct and indirect effect of condition on choice, and (f) studies must be original and experimental. Studies can appear in Meta-Analyses 1 and 2 as long as the relevant dependent variables are reported.

Meta-Analysis 3 inclusion criteria were: (a) studies must manipulate thought listing into a natural and unnatural/reversed condition, (b) participants must do thought listings before making a choice or judgment, (c) the decision task must be an individual-level choice or judgment, (d) studies must report the effect size of framing condition on choice for the natural and unnatural conditions, and (e) studies must be original and experimental. The criteria for Meta-Analysis 1 and Meta-Analysis 3 are mutually exclusive because the former examines unmanipulated query order and the latter examines manipulated (natural versus unnatural) query order. If a paper appeared to meet the inclusion criteria but inadequate data was provided in the manuscript, I referenced supplemental and open access data. If the available materials were not sufficient, I reached out to the corresponding author.

The PRISMA flow diagram (Figure 1.1) illustrates the process of searching and screening papers for inclusion. The literature search was conducted in two waves— a preliminary search in 2020 to scope the literature and design Meta-Analysis 1 and comprehensive database search conducted in July 2024. The database searches return 1,929 records. After removing duplicates and papers that do not mention query theory in the abstract, there were 740 unique papers to screen in detail. Papers were screened three times—once for each meta-analysis (inclusion criteria above).

Each study within a paper was considered independently. For Meta-Analysis 1, 33 effect sizes from 25 papers met the inclusion criteria. For Meta-Analysis 2, 15 effect sizes from 12 papers met the inclusions criteria. And for Meta-Analysis 3, 26 studies from 11 papers met the inclusion criteria.

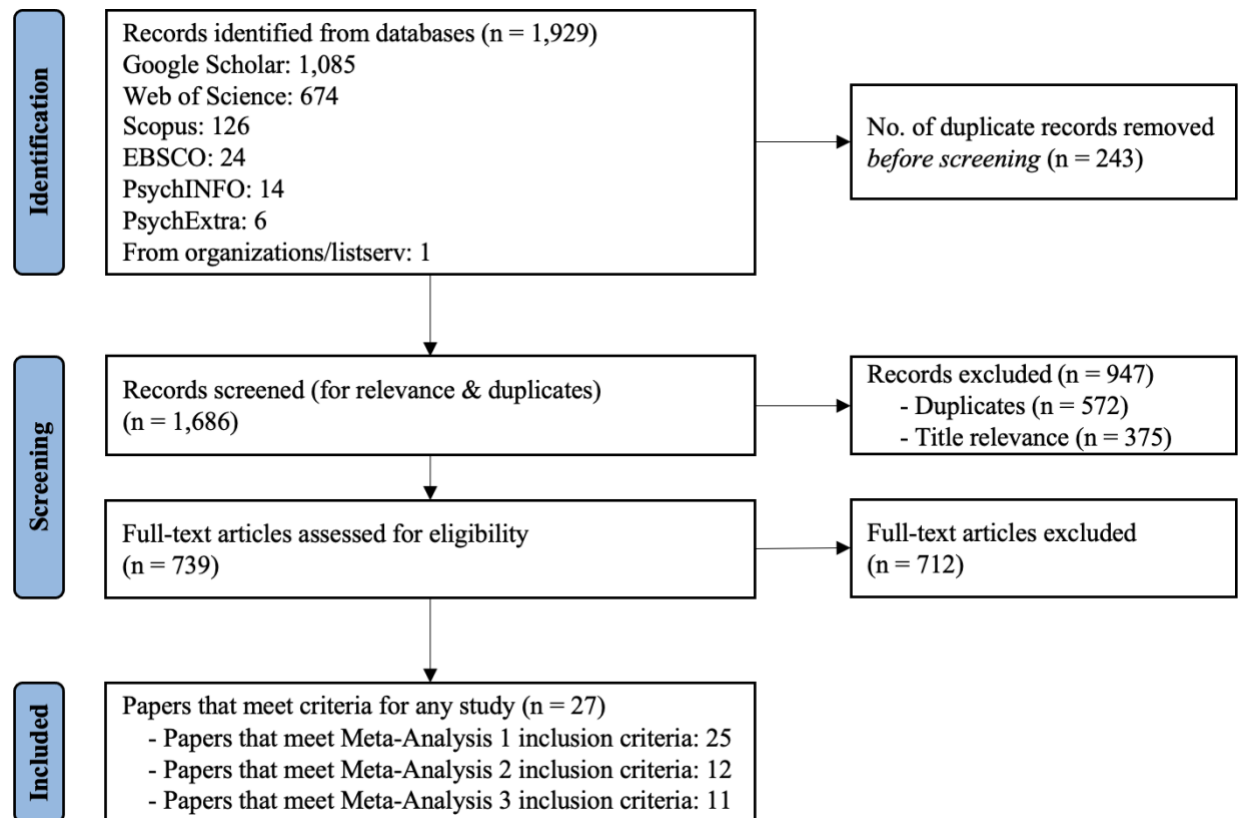


Figure 1.1. PRISMA flow diagram for systematic reviews.

Data extraction, conversion, and coding.

Two of the authors extracted and coded the effect sizes and moderators for the included studies. Disagreements were resolved through discussion with other authors. When manuscripts did not clearly report the necessary metrics or information, I reached out to the corresponding author and discussed the study design.

Meta-Analysis 1. I extracted the effect size for the effect of framing condition on SMRD for all studies that met Meta-Analysis 1 criteria ($k = 33$). The independent variable was the study-specific framing manipulation and the dependent variable was the standardized median rank differences (SMRD) between arguments for choosing Option A and arguments for choosing Option B. SMRD is defined as $2(MR_i - MR_p)/n$, where MR_i is the median rank of arguments for choosing Option A in a respondent's sequence of thought listing preceding the choice, MR_p is the median rank of arguments for choosing Option B, and n is the total number of thoughts listed. SMRD takes the value of +1 if all arguments for choosing Option A are recalled first, and -1 if all arguments for choosing Option B are recalled first, and 0 if the types of arguments do not cluster but are randomly interspersed. If a paper did not calculate SMRD or deviates from the specified equation, I calculated SMRD from the raw data. The outcome variable was always coded so that a higher value is aligned with the decision frame prediction (e.g., in a default frame, the default choice would be considered first). Most studies reported the effect as an F statistic ($k=16$) or a t statistic ($k=13$), with a few effects reported as mean differences ($k=1$) and regression coefficients³ ($k=1$). All effects were transformed to Cohen's d . When a study reported d I used the reported statistic, and when d was not reported I used the 'esc' package in R (Lüdtke, 2016) to convert effect sizes into d to implement the guidelines by Wilson (2023) (updated from Lipsey & Wilson, 2009).

Meta-Analysis 2. For all studies that conduct a mediation analysis of the mediating role of query theory mechanisms (query order, query content, or both), on the effect of condition on

³ The inclusion of regression coefficients in a meta-analysis is sometimes contested because the coefficient describes the relationship between two variables after accounting for other variables that were measured (Roth et al., 2018). I confirmed that the beta extracted from Adjerid et al. (2016) did not include covariates in the regression.

choice, I extracted the proportion mediated, the associated p-value, and the 95% confidence interval. Proportion mediated is the indirect effect divided by the total effect and, simply put, describes how much of the total effect is accounted for by the mediator (MacKinnon et al., 2007). When papers report the proportion mediated, I extract the proportion effect size directly and the p-value and 95% confidence interval from the bootstrapped mediation test (Shrout & Bolger, 2002). When the proportion was not reported, I calculated the proportion from the mediation analysis presented in the paper (i.e., total effect and indirect effect). When raw data was available, I ran the mediation analysis using the “mediation” package in R with 10,000 bootstraps and used the proportion mediated output (Tingley et al., 2014). No effect sizes were converted, the effect size was either extracted directly from the paper, calculated from the direct, indirect, and total effects reported in the paper, or calculated from the raw data. Following best practice for meta-analyses of ratio and proportion metrics, I apply a natural log transformation to the proportion mediated statistic (Higgins et al., 2024). Confidence intervals were not reported for all studies, so I converted p-values to standard error by dividing the natural log of the effect by the z-scored p-value, following the Cochrane Handbook (Higgins et al., 2024).

Meta-Analysis 3a and 3b. Two effect sizes were collected for studies included in these meta-analyses: the effect of the framing condition on the outcome decision when thoughts were listed in the natural order and the effect of the framing condition on the outcome decision when thoughts were listed in the unnatural order. Framing is broadly defined as any intervention that presents information about choice options in normatively irrelevant ways that nonetheless can be expected to change the decision maker’s relative attention to the two choice options, as when the framing changes which of the two choice options is designated as the default choice option. The dependent variable is a decision or judgement; to compare across results, I assign higher values to

align with the decision frame (i.e., the framed option is indicated by higher values). The queries and their order are defined by the context of the study. Natural order is considered the thought listing order (e.g., thought type 1 before thought type 2) that was observed when thought listing was not constrained in any way, Unnatural order reverses that order (e.g., thought type 2 before thought type 1). I extracted the effect sizes and variance for both analyses in Meta-Analysis 3a⁴ (natural query order) Meta-Analysis 3b (unnatural query order)—the independent variable was the decision frame, and the dependent variable was the decision. Studies reported the effects with F statistics ($k=3$), mean differences ($k=3$), t statistics ($k=3$), chi-squared statistics ($k=2$), standardized betas ($k=1$), and Cohen's d ($k=1$). When the effect size was not reported as d in the papers, I convert the reported effect size to d using the 'esc' package in R (Lüdtke, 2016), which follows the Wilson (2023) conversion guidelines (updated from Lipsey & Wilson, 2009).

Moderators. To explore possible sources of heterogeneity in the meta-analyses, I coded for potential moderators. These moderators are characteristics of the observations that could influence differences in the effect sizes or reveal boundary conditions. The moderators are: the year the study was conducted, the publication status (dummy variable, published = 1, not published = 0), the type of choice of each study (risky choice, inter-temporal choice, and multi-attribute choice, each coded as 0 = no, 1 = yes), sample size, and whether the study was conducted online or not (dummy variable, online = 1, in person = 0).

Analysis method

All effect sizes were converted to a common metric (Cohen's d for Meta-Analyses 1 and 3 and proportion mediated for Meta-Analysis 2), as described previously. A random-effects model

⁴ Within each paper, the statistics that I extracted for the natural and unnatural thought order studies were the same (i.e., if a paper compared the effect of decision frame on choice with a t -test, I extracted t -statistics for both the natural and unnatural thought order study).

used the restricted maximum likelihood estimator (Viechtbauer, 2010) for each of four meta-analyses. Analyses with alternative model specifications, including fixed effects and robust variance estimation is presented in Table A1.2 to Table A1.4. Meta-Analysis 1 investigates the relationship between decision frame and SMRD. Meta-Analysis 2 investigates the proportion of the effect of decision frame on choice that is mediated by query theory mechanisms and the model includes fixed effects for the mediator factor that was reported (query order, query content, or a combine factor). Meta-Analysis 3 uses two meta-regressions to investigate the relationship between decision frame and choice in natural query order (Meta-Analysis 3a) and unnatural query order (Meta-Analysis 3b) conditions. Results from the Meta-Analyses 1 and 3 are presented as a Cohen's d (Cohen, 1988) and results from Meta-Analysis 2 are presented as proportion mediated. All analyses are conducted using the 'metafor' R-package (Viechtbauer, 2010).

I tested for heterogeneity by calculating Q and I^2 statistics in the 'metafor' package (Viechtbauer, 2010). Q was calculated as the sum of the squared difference between each included study's effect size and the combined effect size, weighting each study by the inverse variance (Huedo-Medina et al., 2006). A Q hypothesis test is run to examine whether there is true heterogeneity (alternative hypothesis) or not (null hypothesis) in the sample. The I^2 statistic describes how much heterogeneity there is in the sample (Huedo-Medina et al., 2006). Q tests of moderator heterogeneity were used in Meta-Analysis 3 to examine heterogeneity associated with the moderators in those two sub meta-analyses.

When conducting a meta-analysis, it is best practice to account for possible publication bias caused by the systematic underrepresentation of null, small, or counter-intuitive results in academic journals. First, I sought to mitigate this bias in the literature search by including grey literature, such as posters and dissertations, and by contacting researchers who might have

qualifying studies in their file drawer. These methods resulted in the inclusion of eight effect sizes from four papers (these papers are included in the PRISMA flow diagram, Figure 1.1). Second, publication bias is visually assessed with a funnel plot of studies included in the meta-analysis; if no bias exists, the plot is symmetrical and distributed around the combined effect size (Borenstein et al., 2009; Rothstein et al., 2006). Third, I tested for funnel plot asymmetry using Egger's regression test in the 'metafor' R-package (Egger et al., 1997; Viechtbauer, 2010). Finally, when asymmetry was present, I used Duval and Tweedie's (2000) trim and fill method to estimate a symmetrical funnel plot and a bias-corrected combined effect size.

Results

Meta-Analysis 1

Base Model. I found a total of 33 effect sizes from 29 studies and 25 papers, with a total of 11,202 participants, reporting the effect of decision frames on query order. Each effect size was coded for the type of choice—22 effect sizes focus on multi-attribute decisions, 6 focus on intertemporal choice, and 5 focus on risky choice. The median sample size of the studies in this meta-analysis is 189 (SD = 367).

I find a significant effect of decision frame on query order $d^+ = 0.34$, $CI_{95} = [0.27, 0.41]$, $p < .001$, $I^2 = 59.63\%$. This positive effect means that one frame produced a greater standardized median rank difference (SMRD), i.e., an earlier listing of thoughts favoring one choice option than the alternative frame, in the direction predicted by QT, e.g., with defaulted choice options being considered first (Johnson et al., 2007; Weber et al., 2007). The forest plot in Figure 1.2 shows that this effect is positive for all studies in the meta-analysis. The model indicates significant heterogeneity ($Q_E = 82.49$, $df = 32$, $p < .0001$) due to study-level variation. This was further

confirmed by the moderate level of true heterogeneity described by the I^2 statistic ($I^2 = 59.63\%$; Higgins & Thomas, 2023).

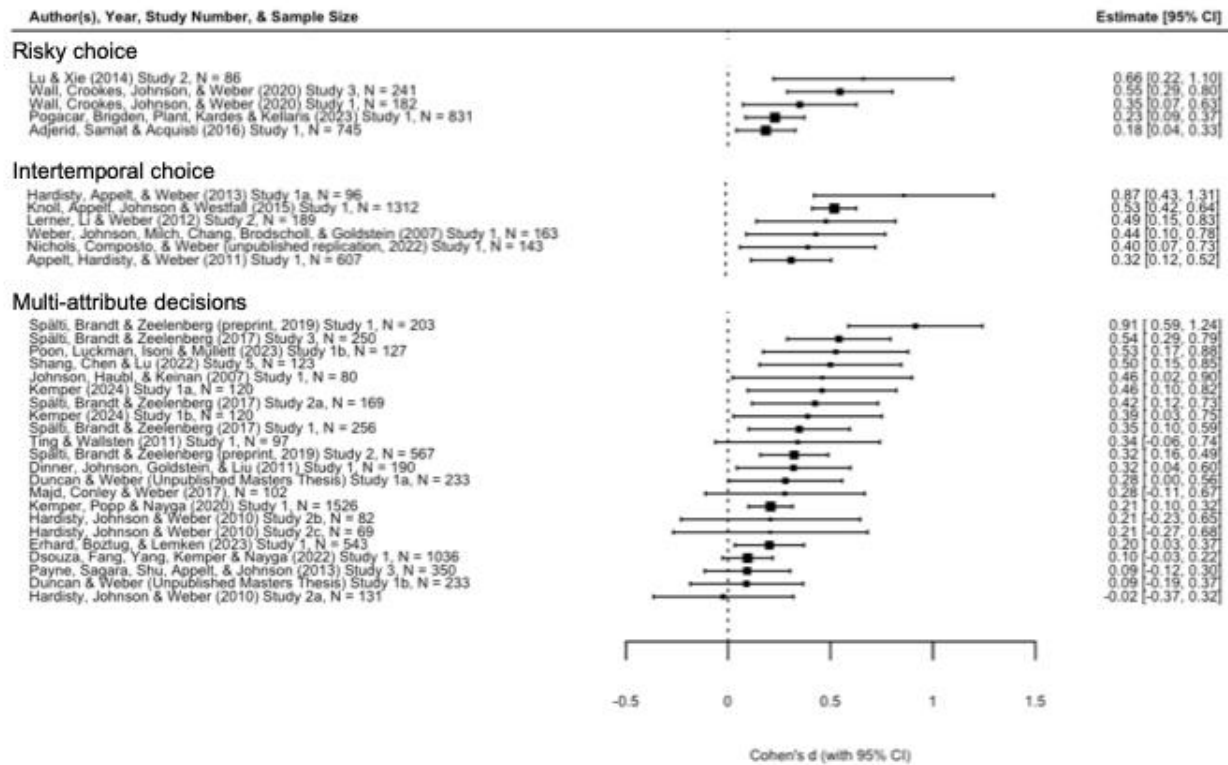


Figure 1.2. Forest plot of 33 effect sizes in Meta-Analysis 1. Effect sizes are measured as the effect of frame on SMRD, using Cohen's d .

The results of study-level ($d^+ = 0.35$, $CI_{95} = [0.28, 0.42]$, $p < .001$, $I^2 = 61.26\%$) and article-level ($d^+ = 0.34$, $CI_{95} = [0.27, 0.41]$, $p < .001$, $I^2 = 59.8\%$) meta-regressions are similar to the base model, showing that the model is robust to observations within the same study (when there are multiple conditions for which effect sizes are provided from a single study) and article (when multiple studies within a paper meet inclusion criteria). These models distinguish between-cluster heterogeneity from within-cluster heterogeneity (Cheung, 2014) because the models contain two

variance components: one for the study or article level (between-cluster) and one for the observations within the studies or articles (within-cluster heterogeneity). I find that the values of both between-cluster and within-cluster heterogeneity are very small (p -value less than .001), thus indicating that most of the heterogeneity in the study occurs due to sampling variance.

Moderators. To explore the possible sources of heterogeneity, I add moderators into the meta-analysis. Each moderator is added and evaluated individually. I evaluated the following moderators: publication year, publication status, if the study focuses on risky choice, if the study focuses on intertemporal choice, if the study focuses on a multi-attribute decision, and sample size. Publication year, publication status, being a risky choice or a multi-attribute decision study, and sample size do not moderate the meta-analysis results; these five moderators have a non-significant effect (all p -values $> .10$; Table 1.1).

Being a study on intertemporal choice, which typically frame either the sooner smaller or the larger later choice options as the default option, significantly moderates the effect of decision frame on SMRD (Table 1.1). Such intertemporal choice studies show a significantly larger effect of frame on the order in which the two choice options are considered in argument listing, compared to other studies. I ran an omnibus test of moderators and found that a significant amount of the model heterogeneity is explained by the intertemporal choice moderator ($F(1, 31) = 4.38, p = .0045$). However, this effect was driven by a single study with a large effect size and, when that study is excluded the moderator was not significant ($p > .10$).

It is still prudent to investigate differences between choice domains to understand the generalizability of the QT mechanism. Therefore, I ran three subgroup analyses for multi-attribute decisions ($k = 22$), intertemporal choices ($k = 6$), and risky choices ($k = 5$). The moderation analyses suggest I should not compare the groups to each other—there is insufficient statistical

power for such an analysis. The meta-analytic effect of frame on query order in multi-attribute decision contexts is $d^+ = 0.30$, $CI95 = [0.22, 0.39]$, $p < .001$, $I^2 = 55.57\%$. The effect in intertemporal choice contexts is $d^+ = 0.48$, $CI95 = [0.33, 0.62]$, $p < .001$, $I^2 = 22.59\%$; and the effect in risky choice contexts is $d^+ = 0.34$, $CI95 = [0.11, 0.56]$, $p = .014$, $I^2 = 61.49\%$.

Moderator	Estimate	SE	<i>t</i> -statistic	<i>p</i> -value	95% Confidence Interval
Year	-0.0011	0.01	-0.15	0.88	[-0.02, 0.01]
Publication Status	-0.0560	0.09	-0.65	0.52	[-0.23, 0.12]
Risky Choice	-0.0045	0.09	-0.05	0.96	[-0.19, 0.18]
Intertemporal Choice	0.1721	0.08	2.09	0.04	[0.00, 0.34]
Multi-Attribute Decision	-0.1062	0.07	-1.51	0.14	[-0.25, 0.04]
Total N	-0.0001	0.00	-1.53	0.14	[0.00, 0.00]

Table 1.1. Moderator effects for Meta-Analysis 1. Each moderator is added, individually, to the main meta-regression and the estimate is reported.

Publication Bias. To look for possible publication bias, I created a funnel plot to display each Cohen's d as a function of the standard error (Figure 1.3a). A number of observations appear outside of the funnel on both sides, which could suggest potential biases (i.e., studies missing from the literature due to suppression of extreme results, poor methodology, etc.). I ran Egger's test (Egger et al., 1997) for funnel plot asymmetry and find some evidence of asymmetry, $t(31) = 2.25$, $p = 0.03$.

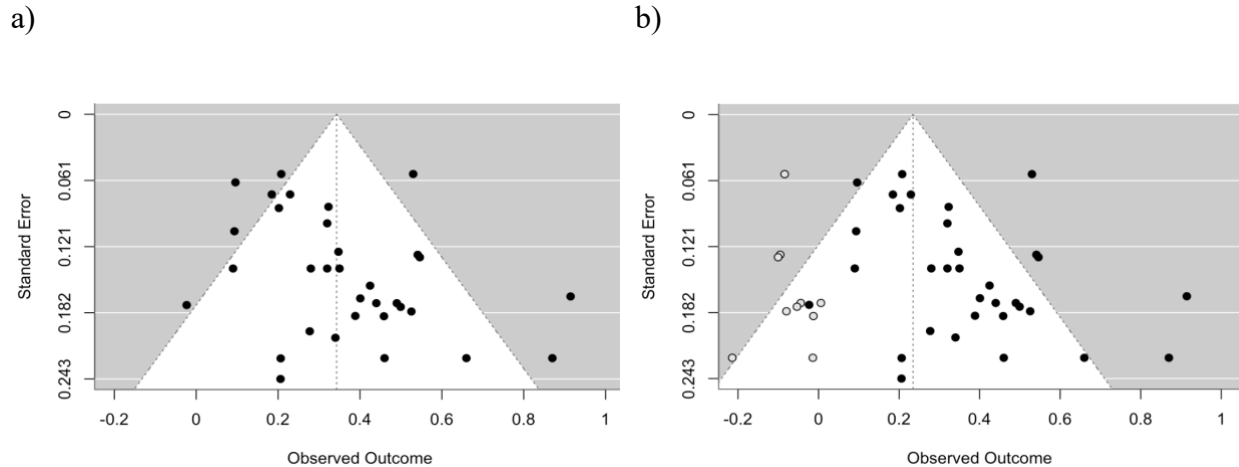


Figure 1.3. a) Funnel Plot and b) Trim & Fill Funnel Plot for Meta-Analysis 1.

To test the sensitivity of the model to different types of biases, I use Duval and Tweedie's (Duval & Tweedie, 2000) trim and fill method. This model estimates the number of missing studies due to suppression of results for reasons such as publication bias or extreme results. It does so by adding synthetic effect sizes to the observed data to create a symmetric funnel plot. It is important to emphasize that this method does not make a more valid estimate, but simply tests the sensitivity of the data to different selection mechanisms. The trim and fill model included 12 missing studies, which make the overall estimated effect $d^+ = 0.23$ (CI95 = [0.15, 0.31], $p < .001$). This result indicates that the direction and magnitude of the effect is robust to suppression biases.

Meta-Analysis 2

To understand the extent to which the QT mechanisms mediate and thus explain the effect of decision frame on choice, I identify 15 studies from 12 papers with data from 2,535 participants. These mediation analyses are summarized in Table A1.1. Of the 15 mediated proportions, nine are for a SMRD mediator, three are for a proportion of thoughts mediator, and three are for a combined factor mediator (SMRD and proportion of thoughts combined in some way). Ten of the 15

mediation effects collected are significant, indicating that the QT mechanisms do account for some proportion of the effect of decision frame on choice. All six of the discounting studies find significant mediation (Table A1.1). This descriptive analysis suggests that QT processes are one of the contributing mechanisms of framing effects. As illustrated by the five unmediated studies, there are other mechanisms that also explain these effects.

Using a linear meta-regression with fixed effects for the type of mediator (query order, query content, or a combined factor), I find that the overall proportion mediated is 0.55 (CI95 = [0.31, 0.98], $p = 0.04$, $I^2 = 25.2\%$). There is no significant effect of the mediator type on the proportion mediated effect estimate. Examining the Q test for heterogeneity, I did not find significant heterogeneity ($Q_E = 10.39$, $df = 12$, $p = 0.58$). I examined the omnibus test of moderators to examine whether the type of mediator affected the combined effect, and I do not find a significant impact ($Q_M = 1.93$, $df = 2$, $p = 0.38$). I do not find evidence of publication bias in this study, measured by Egger's test ($p = 0.08$). Aligning with the recommendations on how to meta-analyze proportion measures, I use a natural log transformation to analyze proportions (Higgins et al., 2024).

Some scholars argue that proportion mediated should not be calculated because the metric is very unstable—particularly when the total effect in the mediation analysis is small (MacKinnon et al., 1995). I conduct a leave-one-out sensitivity analysis to assess the influence of individual studies on the overall effect size. The results indicated that the pooled estimate remained stable across all iterations, such that the confidence intervals always contained the overall meta-analytic effect of 0.55 (Table A1.5). This suggests no single study unduly influences the meta-analytic findings. Calculating standard error from a bootstrapped mediation analysis p -value is imprecise when exact p -values are not reported, as is the case for some of the effect sizes. To address this

concern, I evaluate the overall proportion mediated by calculating the pooled mean. This strategy does not rely on standard error and simply weights the different effect sizes (proportion mediated) by the sample size—the pooled mean mediation proportion is 27%.

Meta-Analysis 3

To investigate the causal role of query order, I meta-analyze studies that use an explicit order manipulation of thought listings. The effect size of frame condition on decision outcome is meta-analyzed for the natural order and opposite or unnatural order separately. The size of frame on decision when thoughts are explicitly listed in the natural query order is $d^+ = 0.92$, $CI95 = [0.74, 1.09]$, $p < .001$, $I^2 = 49.24\%$. This large positive effect indicates that requested thought listing in the natural querying order results in a large difference in choice under the two frames. I find a much smaller effect of frame on choice in the unnatural query order condition, $d^+ = 0.39$, $CI95 = [0.25, 0.53]$, $p < .001$, $I^2 = 27.18\%$.

The difference between these two effect sizes is evidence for the causal effect of query order on choice. I compared the effect size differences between the natural and unnatural studies in each paper by calculating the precision-weighted difference to account for standard error variability between studies. The precision-weighted difference was 0.55 ($SE = 0.08$; $CI95[0.40, 0.70]$). This is supported by the forest plot in Figure 1.4, which showed that average effect of frame on decision is much larger under the natural query order condition than under the unnatural query order in almost all studies.

There was significant heterogeneity in the natural order model ($Q_E = 23.77$, $df = 12$, $p = 0.02$), but not in the unnatural order model ($Q_E = 16.15$, $df = 12$, $p = 0.18$). The level of true heterogeneity from study-level variation in the natural order model was moderate ($I^2 = 49.24\%$). Examining the publication bias in these two meta-analyses, I found that neither meta-analysis has

significant funnel plot asymmetry, as measured by Egger's test (natural query order: $t(11) = -1.76$, $p = 0.11$; unnatural query order: $t(11) = -0.88$, $p = 0.4$).

Next, I examined the choice domain as a potential moderator. In each sample, 6 studies focus on a multi-attribute decision, 4 focus on intertemporal choice, and 3 focus on risky choice. Topic was not a moderator to the natural order and unnatural order meta-analyses. Using an omnibus test of moderators, I found that heterogeneity in the effect size is not explained by choice domain in the natural order meta-analysis ($F(2,10) = 2.28$, $p = .15$) nor in the unnatural order meta-analysis ($F(2,10) = 0.99$, $p = .41$). These results should be interpreted with caution due to the small number of studies included in the sample ($k=13$).

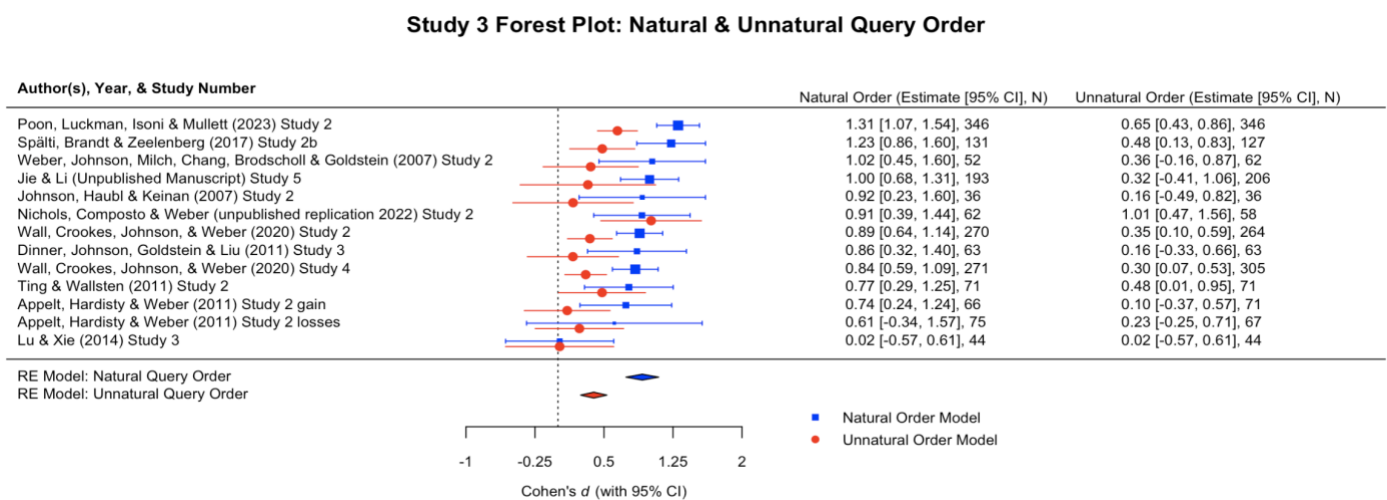


Figure 1.4. Forest plot of 13 effect sizes in natural order studies and 13 effect sizes in unnatural order studies from Meta-Analysis 3. Effect sizes are measured as the effect of frame on choice, using Cohen's d , in either a natural thought order or unnatural thought order condition. Each study has two effect sizes, one in the natural thought order condition and one in the unnatural thought order condition; effect sizes from the same study are shown on the same line. Effect sizes from

natural order conditions are shown in blue and effect sizes from unnatural order conditions are shown in red.

Discussion.

Research on QT over the last seventeen years supports the theory's attention and memory retrieval process account of framing effects and the general notion that preferences are constructed rather than economic primitives (Weber & Johnson, 2006). In particular, QT provides a convincing set of psychological mechanism that give rise to preferences that are well modeled by PT's loss-averse value function. I systematically reviewed all 27 papers that have used a QT paradigm and investigated QT's process theory claims. For the effect of frame on query order, a total of 33 effect sizes ($n = 11,202$) were found, with 22 ($n = 4,466$) in the domain of multi-attribute decisions, 6 ($n = 1,218$) in the domain of intertemporal choice, and 5 ($n = 1,015$) in the domain of risky choice. The results from the random-effects model meta-analysis suggest that decision frame has a significant small effect on query order (measured as SMRD), after accounting for possible publication bias ($d^+ = 0.23$, $CI95 = [0.15, 0.31]$, $p < .001$). However, there is significant heterogeneity in this model. Through moderation analysis, I found that studies in the intertemporal choice domain have significantly larger effect sizes of frame on query order than those in the two other domains. The meta-analytic effect is stable over the other moderators I investigated—publication status, publication year, sample size, being in the risky choice domain, and being in the multi-attribute decision domain. These findings align with the QT premise that decisions are deconstructed into component queries (Collins & Michalski, 1989) and that query order is affected by decision frame (Johnson et al., 2007; Weber et al., 2007). Taken together, Meta-Analysis 1

provides evidence that framing affects decisions by shifting attention from one response option to the other, with the two options being evaluated in different orders.

To explore the causal claims in QT I conducted two additional meta-analyses. Meta-Analysis 2 investigated the claim that query order and/or the associated balance of evidence for the two choice options moderate the effect of decision frame on choice. I collected 15 effect sizes ($n = 2,535$) from mediation analyses in QT studies to explore how much of the effect of decision frame on choice is mediated by query order and/or query content and find convincing evidence that the QT mechanisms mediate framing effects. Meta-Analysis 3 investigated the claim that explicitly interfering with the natural shift in attention by reversing the natural query order in the two framing conditions will reduce framing effects (Johnson et al., 2007; Weber et al., 2007). If attention-mediated automatically initiated differences in query orders under the two frames do indeed (at least partially) explain the framing effect, then explicitly requesting argument generation in the natural order should increase the framing effect and explicitly requesting argument generation in the opposite unnatural query order should reduce or eliminate the effect. I collected 16 effect sizes ($n = 3,401$) for Meta-Analysis 3—13 ($n = 1,681$) with the natural query order and 13 ($n = 1,720$) with the unnatural query order, and found that the meta-analytic effect of frame on choice in natural order studies is $d^+ = 0.92$, ($CI_{95} = [0.74, 1.09]$, $p < .001$) and in unnatural order studies is $d^+ = 0.39$ ($CI_{95} = [0.25, 0.53]$, $p < .001$). The 95% confidence intervals do not overlap, suggesting that the framing effect significantly decreased when query order was explicitly manipulated from its natural order to the opposite order. Together, Meta-Analyses 2 and 3 support the causal role of the QT mechanisms in choice across the three choice domains of risky, intertemporal, and multi-attribute decisions.

This is the first paper to meta-analytically evaluate the claims of QT and to examine its psychological process theory across studies and contexts. There are limitations to acknowledge. First, all studies use the same thought-listing paradigm which could limit the generalizability of these findings. The QT typed concurrent report method is more effortful than a spoken report process tracing, I suspect that these reports are a less complete representation. This paradigm has been validated, and many studies have produced choice results that are very close to those reported in other studies that do not use typed reports. In particular, the QT paradigm was validated using an implicit association task to examine whether the thought listing task affected the memory retrieval process (Weber et al., 2007). The content clustering (i.e., query order) measurement in the QT paradigm was also validated using latent semantic analysis which confirmed participant's thought categorizations (Johnson et al., 2007).

Additional limitations include the stability of the proportion mediated metric and the number of studies for choice types. There are critiques of the stability of proportion mediated (MacKinnon et al., 1995), which is the dependent variable in Meta-Analysis 2, and the mediator is not consistent across all studies. I discuss these critiques and take two precautions to avoid over-interpreting these results—first, I collect and report the type of mediator used, and second, I produce a summary table and discuss the results qualitatively (Table A1.1). Finally, the number of studies in the risky choice and intertemporal choice domains is relatively small ($k = 5$ and $k = 6$, respectively) and the results of the meta-analysis should be interpreted with caution (Steel & Kammeyer-Mueller, 2002; Thompson & Higgins, 2002).

In conclusion, this set of meta-analyses supports the QT process theory of framing affects. There is strong evidence that query order explains framing effects in the contexts of risky choice, intertemporal choice, and multi-attribute decisions. Although there are fewer studies examining

the causality of query order, I find convergent evidence that framing effects drive anomalous decision making by affecting query order. Future research should aim to expand QT processes to other the choice contexts, such as multi-alternative decisions (Poon et al., 2023) and broaden the types of framing devices (such as social information and contextual signals).

Chapter 2: A memory-retrieval perspective on the construction of social norm perceptions

The previous chapter explored the construction of individual preference through memory-retrieval processes by presenting a meta-analysis of query theory. In the present chapter, I propose and test an expansion of query theory to the domain of social cognition, specifically the perception of social norms. This investigation is motivated by the growing literature documenting pluralistic ignorance on support for environmental policies and pro-environmental behaviors (Mildenberger & Tingley, 2019; Geiger & Swim, 2016), leading to false social realities (Sparkman et al., 2022). People underestimate the general public's support for pro-environmental policy (Sparkman et al., 2022) and attitudes (Geiger & Swim, 2016); in other words, the perceived social norm is systematically lower than the true norm on these topics. This pattern of underestimation leads to self-silencing and is a barrier to individual and collective action in domains such as climate change (Geiger & Swim, 2016). The literature has focused on characterizing these gaps and their consequences, with less focus paid to the drivers and underlying cognitive processes of these inaccurate social norm perceptions. This chapter reviews existing theories of norm perception in the literature and offers a memory-retrieval theory to explore the cognitive mechanisms of social norm perceptions. I explore the role of memories in norm perceptions derived from the combined intuition of Kahneman and Miller's norm theory (1986) and Weber and Johnson's query theory (Johnson et al., 2007; Weber et al., 2007). Methodologically, this chapter leverages query theory's thought-tracing paradigm while participants evaluate different norm scenarios. To develop and test the query theory explanation of social cognition, this chapter surfaces and characterizes the natural thought process that people go through when estimating the prevalence of a social norm.

Origins of social norm perceptions

Social norm perceptions are shaped by the public behavior of others, which helps individuals understand and adhere to societal expectations. These perceptions may well be biased and will be influenced by features of the perceiving individual, of the social group or norm reference, and the environment. Social norms evolve through social interactions and cultural transmission, influenced by ecological and historical factors (Gelfand et al., 2024). This dynamic process can lead to the persistence, change, or erosion of norms over time. Observing or learning about others' behavior can inform both the set of permissible and expected behaviors in a given context (Miller & Prentice, 1993; Cialdini, Reno, & Kallgren, 1990). In Cialdini and colleague's classic study on littering, the authors found that modeled behavior (i.e., littering in a clean vs. heavily-littered park) informed the observer's descriptive norm perception, while the reaction to that behavior (i.e., cleaning up the litter vs. leaving the litter) shaped injunctive norm perception (Cialdini, Reno, & Kallgren, 1990).

These perceived norms can be inherently biased by the differences between public and private behaviors. For example, students on college campuses have inaccurate norm perceptions about the average amount of binge drinking that happens on college campuses (Miller & Prentice, 1993). This overestimation is driven by salient memories of public behavior which outweigh consideration of private, non-binge drinking behaviors. Building on this line of work, Paluck and Shepherd found that not all peer signals are equal; norm signals from a peer that received more attention were more impactful than signals from unpopular peers (Paluck & Shepherd 2012). Social norm perceptions are shaped by the interplay of individual, group, and environmental factors and moderated by whether a behavior is public and how social capital of the messenger.

Individuals acquire information about social norms through direct interaction with or observation of others, through summary information transmitted by others, and through institutional signals (Tankard & Paluck, 2016; Lapinski & Rimal, 2005). Summary information about a group, also termed social norms marketing, involves the communication of common behavior (descriptive norms) or attitudes (injunctive norms), often through media, signage, or advertising. In a pair of field experiments, a descriptive social norm communication (e.g., “the majority of guests reuse their towels”) increased conformance to the targeted pro-environmental behavior (Goldstein et al., 2008). This style of intervention has received much attention from psychologists, particularly in the study of environmental behavior and attitudes (Nolan, 2021). A meta-analysis of the literature found a significant though small effect of social norm interventions on attitudes, norm perceptions, and behaviors, but also some conflicting results, with evidence of backfire and partisan effects (Nolan, 2021; Rhodes et al., 2020). For example, Goldstein’s hotel towel study failed to replicate in a German sample, which the authors attribute to the cultural differences between German and American environmental attitudes (Bohner & Schlüter, 2014). Partisan effects in this domain appear as the systematic underestimation of Republican concern about climate change and the willingness of each political party to work with the other on climate policy (Van Boven et al., 2018).

In the field, news media coverage is a major source of summary norm information. Media coverage can amplify this information explicitly or implicitly through the cultural understanding that mass media covers popular topics (Lazarsfeld & Merton, 1948). Media also amplifies institutional signals, such as rulings from the Supreme Court; studies found Court rulings affect norm perceptions, in part through media conversation and reaction (Tankard & Paluck, 2017; Clark et al., 2023). Media coverage amplifies information through simultaneous co-attention, as the

viewer is aware that everyone is being exposed to the same information (Shteynberg et al., 2016). All these signals, whether directly observed, summarized, or directed by institutions, contribute to the evidence base from which an individual's social norm perception emerges or is activated. These perceptions are shaped by cognitive processes and cultural dynamics, often leading to biases and misperceptions.

The memory-retrieval perspective

Kahneman and Miller's norm theory (1986) proposed that category norms are constructed through ad hoc recruitment of specific and exemplar representations. The myriad norm signals that people encounter throughout life are not aggregated and stored as norms but rather the norms are constructed through "backward thinking" (p. 137). Norm theory does not claim ad hoc construction is the only mechanism and suggests anticipation may also inform perceptions. The focus on backward process leverages the exemplar model, which suggests that several representations are evoked at once by a stimulus. For example, being asked to think about corporate environmental responsibility (the stimulus) would cue thoughts about the typical and experienced function of corporations, climate friendly products and initiatives, environmental marketing, pollution and oil spills, and maybe the way others have discussed or reacted to those products and pollution events. Availability of exemplars increases perceived normality, i.e., a stimulus that elicits many examples of a given behavior, even if they are not representative, will be considered more normative than a stimulus that elicits fewer exemplars. If the corporate environmental responsibility stimulus cues many examples of sustainable products that the participant has seen others use, it might be considered more normative that companies produce sustainable products. The availability heuristic is influenced by recency, frequency, intensity, area, duration, meaningfulness, familiarity, and self-involvement—a non-exhaustive list that are not mutually exclusive (Avant & Helson, 1973). The

norm theory perspective contrasts with the traditional view that events, or stimuli, are interpreted using stored schemas and frames of reference. The memory-based construction of social norms makes norm perceptions dynamic and context-dependent. Therefore, norm perceptions are highly mutable to the availability heuristic.

The role of affect serves as an amplifier of exemplars, with particularly surprising examples drawing increased attention. This phenomenon can be understood through several psychological mechanisms. First, the 'change of standard' effect (Higgins & Lurie, 1983) demonstrates how past events can be reconstructed in contemporary contexts, altering their affective evaluation. Second, social and environmental features can further enhance memory encoding. For instance, mass media communication increases processing through co-attention, which may increase weight (Shteynberg et al., 2016).

The concept of constructed norms suggests that the exemplars that come to mind are heavily influenced by stimulus characteristics. In the context of food choices, this means that a stimulus signaling vegetarian meals will more readily elicit memories about vegetarian meals or practices than a stimulus that highlights meat-based meals would elicit. Kahneman and Miller (1986) argue that people tend to over-generalize from initial exemplars when estimating category characteristics. For example, observing a single member of an unfamiliar group choose a meat dish at a restaurant might lead the observer to conclude that eating meat is typical for the entire group. This process of exemplar retrieval operates through a concentric search, where the initial exemplar shapes subsequent exemplars. Critically, this suggests that norm perceptions do not regress to the mean (Kahneman & Tversky, 1973; Ross & Anderson, 1982), but instead can be skewed by initial, potentially unrepresentative examples. Returning to the food choice example, a stimulus that

frames the question of normative meal choice around vegetarian meals would first elicit a memory about a vegetarian meal, which could cascade through subsequent exemplar search.

Query theory offers a specific process explanation of how norms are constructed through memory retrieval, searching for and evaluating the exemplars that are cued by a stimulus (i.e., the experimental manipulation). Query theory suggests that features of the memory retrieval process affect preferences and can change individual choice (Johnson et al., 2007; Weber et al., 2007). First, theory on plausible reasoning suggests that people break down a valuation question into component queries (Collins & Michalski, 1989). For example, evaluating normative meal choice may be deconstructed into: “why do people choose vegetarian meals?” and “why do people choose meat meals?” Second, individuals access more information from memories closely associated to the question topic or stimulus (Higgins, 1996). Empirical evidence shows that the semantic frame of a question affects the contents of a speaker’s reasoning (McKenzie & Nelson, 2003). So, asking whether vegetarian meals are normative for a group would direct attention to features of the vegetarian meal. And third, memory interference and inhibition play an important role, where the recall of memories related to the first query suppresses access to memories related to later queries (Dempster, 1995; Anderson, Bjork, & Bjork, 1994). When thinking about reasons a vegetarian diet is or is not a common for a group, memories related to meat dishes are suppressed.

These memory processes explain how frame affects individual preference, such as a perspective frame (e.g., buyer versus seller role in an endowment paradigm; Johnson et al., 2007) or a temporal reference frame (e.g., accelerating versus delaying reward in an intertemporal choice paradigm; Weber et al., 2007). The literature on individual choice also finds that this set of processes leads to ‘wrong’ or ‘irrational’ choices for two reasons: (a) the memory base on which perceptions are formed may not be representative, and/or (b) memory recall may be systematically

biased. The present chapter investigates whether such memory-retrieval processes affect the perception of social norms.

This chapter offers a quantitative exploration of norm construction through memory-based cognitive process tracing. Memories shape who we are, what we believe, and how we interact with the world. As articulated in Kahneman and Miller's (1986) norm theory, memories hold information, broadly defined, about the world and about others that can shape social impressions and normativity. When probed with a category stimulus, i.e., a clear reference to a domain such as corporate behavior or meal choice, element recruitment is focused on exemplars from within that category domain (Rips & Turnbull, 1980). The query theory paradigm applied in the present studies tests the theory of constructed norm perceptions and examines the effect of different frames (i.e., highlighting one options or the other in the stimulus) on both the memory-retrieval process and the resulting perception of the strength of the norm.

Episodic and semantic memory

Tulving's theory of memory (1972; 1985) proposes two distinct yet interconnected forms of memory—*semantic memory*, the memory of general knowledge and language, and *episodic memory*, memories that are associated with a specific time and place. Semantic memory is organized around concepts and is informed by comprehension, concepts, and ideas. Episodic memory is organized around temporality and is informed by sensations, events, and episodes (Tulving, 1983). Episodic recall is the ability to reconstruct both the content and the context of acquisition. Context is relevant in distinguishing when and to whom an action or attitude is normative (Tulving, 2002). In the present studies, I investigate the extent to which semantic or episodic memories are retrieved when participants construct social norm perceptions. I examine

whether the type of norm that is constructed (injunctive versus descriptive) will activate different types of memories (semantic versus episodic memories).

Social norm perceptions and pro-environmental behavior

Social norm perceptions shape pro-environmental behavior and attitudes. Social norm perceptions set expectations about how people do and should act in specific context, influencing a range of individual and collective behavior, such as electricity usage (Allcott, 2011; Costa & Kahn, 2013; Schultz et al., 2007), recycling (Andersson & von Borgstede, 2010) and household water use (Ferraro et al., 2011). This literature often explains social norm interventions as the effect of peer influence and conformity. Although much of the literature on social norms and environmental behavior has focused on low-impact individual behaviors (for discussion, see Nielsen et al., 2021), changing norms and norm perceptions can lead to and reinforce critical social tipping points with large ecological and social impacts (Castilla-Rho et al., 2017; Otto et al., 2020). Norms can also act as a signal to institutions (corporations or government) about what the public expects or considers permissible practices (Otto et al., 2020). The present chapter uses norms from two domains to capture these two pathways, institutional signals and peer influence, by which norms affect systemic change. In Study 1 I examine norm perceptions related to corporate priorities and practices that impact the environment and in Study 2 I examine norm perceptions about meal choice in a public setting.

Norm perceptions about corporate practices. The model of social tipping for stabilizing Earth's Climate put forward by Otto and colleagues (2020) argues that the customs, norms, and values of society shape financial markets and resource allocation. The authors suggest multiple pathways for this influence—both direct through public support or opposition for changes in the fossil fuel industry (i.e., divestment, building new plants) or indirectly through policy (i.e., funding

for new technology development, improving environmental standards). In Study 1, I measure norm perceptions about corporate economic priorities and environmental practices that affect the environment. Each norm summarizes a common business strategy—prioritizing the bottom line or engaging in practices that support environmental and social outcomes. The first references the traditional view of fiduciary responsibility to the shareholder and the second references the contemporary interpretation of fiduciary duty to achieve environmental, social, and governance outcomes (McGowan, 2022).

Norm perceptions about meal choice. Food and meals are deeply embedded in social and cultural contexts; it is broadly acknowledged that what we eat is shaped by others (Higgs, 2016). Reducing meat consumption is one of the most impactful individual actions to improve planetary health because the production of meat, particularly red meat, is resource intensive and threatens the planetary boundaries (Willett et al., 2019; Campbell et al. 2017; Ivanova et al., 2020). Average meat consumption also exceeds physician-recommended levels and has adverse health impacts (Westhoek et al., 2014). Meat consumption behavior is reinforced publicly through social and environmental cues (meat options are the typical default in restaurants) and privately through habits (the frequent preparation of meat-based meals). There is evidence that perceived social norms can reduce meat consumptions; an experiment by Sparkman and Walton (2017) found that a dynamic descriptive norm intervention reduced meat consumption in both lab and field contexts. A recent cross-sectional survey of U.K residents found that context (i.e., supermarket versus restaurant versus work cafeteria) elicited different descriptive and injunctive norm perceptions; the strongest meat consumption norms, both descriptive and injunctive, were reported in the supermarket (Wolfswinkel et al., 2024). The authors also found that personal norms, i.e., the amount of meat the participant ate, predicted their norm perceptions—people who ate more meat

regularly perceived a higher norm of meat consumption compared to those who ate less meat. Taken together, these findings that norm perceptions vary by individual and contextual features suggest that they are constructed by memory retrieval processes that may reveal similar variation.

Chapter Overview

This chapter seeks to (1) investigate whether social norm perceptions are constructed through memory-retrieval, (2) identify the features of memory that are elicited when norm perceptions are constructed, and (3) evaluate norm perceptions about two environmental topics—corporate responsibility and individual food choice. The following studies explore whether social norm perceptions are constructed at least in part through memory-retrieval, and the extent to which norm perceptions are affected by the order and content of those memories. Study 1 was a first attempt to use the query theory mechanism to study social cognition and investigated whether question frame and queried norm type affect the reported strength of a social and an antisocial norm. Study 2 investigated the same question about frame and norm type in a different norm domain context that was more familiar to the average participant.

Study 1 was conducted as an exploratory analysis to investigate how norm perceptions might be explained by the query theory mechanism of memory-based construction. Specifically, I evaluated the effect of question frame (agree vs. disagree) and norm type (injunctive vs. descriptive) on query order and content. The prominence of episodic and semantic memories was also examined using *post-hoc* classification of the thought listings.

Study 2 formally tested the query theory explanation of norm perception: *Hypothesis 1 (H1)*: Question frame affects query order (H1a) and content (H1b), such that the framed option (i.e., the outcome that is cued in the stimulus) versus the unframed option is attended to first and

receives a higher proportion of supporting (rather than opposing) thoughts. *Hypothesis 2 (H2)*: Norm type (injunctive vs. descriptive) affects query order (H2a) and content (H2b); this hypothesis does not have a direction. *Hypothesis 3 (H3)*: Query order (H3a) and content (H3b) predict norm perception, such that the option attended to first (query order) and more (query content) is perceived to be a stronger norm. Memory coding was also formalized in Study 2, so that I could conduct an exploratory analysis of the types of memories (semantic versus episodic) that occur in the different norm type conditions (injunctive vs. descriptive).

Study 1: CCS norm estimates

Method

Study 1 provided a first investigation of the thought processes elicited by norm evaluation. Participants read a statement that described possible normative behavior from companies working on energy projects (e.g., about economic priorities: “Energy companies prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community.” or about environmental practices: “Energy companies adopt practices to reduce the environmental impact of their operations, even when it is costly or risky to them.”; Table 2.1). Using a 2x2 between-subject experimental design, participants are randomly assigned to either an agree or disagree frame (“frame”) and either an injunctive or descriptive norm perception (“norm type”). To improve the generalizability of the study, I used two norm topics, one about economic priorities and the other about environmental practices with the stimulus text provided in Table 2.1. After reading the stimulus text, participants were asked to share all thoughts that went through their mind as they considered whether the described practices or priorities were a norm that existed in the world. After that, they answered the main dependent variable about norm perceptions, namely

whether most people would (dis)agree with the statement. Finally, participants were re-shown each thought they had listed and asked to categorize each as “norm-increasing” (a thought about why/if the statement is a norm), “norm-decreasing” (a thought about why/if the statement is not a norm), both or neither. A separate control group reported their norm perceptions on a randomly selected two of the four topic x norm-type combinations without any thought-listing instructions. Study 1 was exploratory; the sample and procedure are preregistered as part of a separate project (As.Predicted #165524).

Topic	Frame	Norm Type	Stimuli Text
Economic Priorities	Agree	Injunctive	Read the statement below and consider whether <u>this is a commonly held norm</u> , i.e., whether most people agree with the statement: <i>Energy companies <u>should</u> prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community.</i>
Economic Priorities	Disagree	Injunctive	Read the statement below and consider whether <u>this is not a commonly held norm</u> , i.e., whether most people disagree with the statement: <i>Energy companies <u>should</u> prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community.</i>
Economic Priorities	Agree	Descriptive	Read the statement below and consider whether <u>this is a commonly held norm</u> , i.e., whether most people agree with the statement: <i>Energy companies prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community.</i>
Economic Priorities	Disagree	Descriptive	Read the statement below and consider whether <u>this is not a commonly held norm</u> , i.e., whether most people disagree with the statement: <i>Energy companies prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community.</i>
Environmental Practices	Agree	Injunctive	Read the statement below and consider whether <u>this is a commonly held norm</u> , i.e., whether most people agree with the statement:

			<i>Energy companies <u>should</u> adopt practices to reduce the environmental impact of their operations, even when it is costly or risky to them.</i>
Environmental Practices	Disagree	Injunctive	Read the statement below and consider whether <u>this is not a commonly held norm</u> , i.e., whether most people disagree with the statement: <i>Energy companies <u>should</u> adopt practices to reduce the environmental impact of their operations, even when it is costly or risky to them.</i>
Environmental Practices	Agree	Descriptive	Read the statement below and consider whether <u>this is a commonly held norm</u> , i.e., whether most people agree with the statement: <i>Energy companies <u>are adopting</u> practices to reduce the environmental impact of their operations, even when it is costly or risky to them.</i>
Environmental Practices	Disagree	Descriptive	Read the statement below and consider whether <u>this is not a commonly held norm</u> , i.e., whether most people disagree with the statement: <i>Energy companies <u>are adopting</u> practices to reduce the environmental impact of their operations, even when it is costly or risky to them.</i>

Table 2.1. Stimuli detail from Study 1.

Participants

I recruited 3,771 participants to this study. This study was embedded in a larger survey about energy technology that was deployed to a representative sample of the United States public ($n = 1,945$) and a location-targeted sample based on proximity to energy projects ($n = 1,826$). Study 1 was pre-registered (As.Predicted #165524). Both samples were recruited by a third-party survey panel provider, Centiment, in February 2024. The representative public sample was quota-matched to be nationally representative by income, geography, race, gender, age, education, and political party. The geographic distribution matched the regions in the United States Census: 17% of responses were collected from the Northeast, 21% from the Midwest, 38% from the South, and 24% from the West. The location-targeted sample was recruited from ZIP codes located near a

carbon capture and storage project site that was operational or under development at the time of the survey⁵.

As preregistered, participants who failed the attention check were excluded from analysis, resulting in 16 exclusions. Four participants in the treatment (thought listing) conditions did not list any thoughts and were also excluded. This left us with 3,751 participants. The mean age across both samples was 47 years, with a standard deviation of 17 years (Table A2.1). 57% of all samples identified as female. The median annual household income bracket was \$20,000-40,000, and 36% of respondents had received a college degree. The distribution of political affiliation was 30% Democrat, 28% Independent, 31% Republican, and 11% other. Full demographic information is presented in Tables A2.1.

A post hoc⁶ power analysis was conducted to evaluate the statistical power of the study's main analyses. First, I examined the power for a t-test of the effects of frame on norm perceptions. These analyses are run separately on each of the four subgroups (norm type x topic), e.g., the effect of frame on the norm perceptions of the injunctive environmental practices norm is analyzed on its own. The relevant results are shown in Table 2.3. The average sample size per group in our study is 245 and the average effect size is 0.19. When conducting a two-sample t-test with a two-

⁵ Project sites for the location-targeted sample were selected from the US Carbon Capture Activity and Project Map from the Clean Air Task Force and Global CSS Map from the Scottish Carbon Capture & Storage research group. I included operational, under construction, and planned projects involved in direct air capture, sequestration, and reutilization in fuel manufacturing and excluded the minority of projects using carbon capture in cement manufacturing and power plants. To ensure planned and under development projects were visible to the public, I conducted a media search for announcements of CCS projects in local newspapers through January 2024. Projects were categorized as including capture, transport, and/or storage elements based on information from these aforementioned sources as well as company announcements. For each project, I identified the ten ZIP codes closest to the location of the CSS project site, resulting in a sample of 21 CCS project sites with a combined local population of over 1.7 million residents across nine states.

⁶ The present study was embedded in a larger survey about perceived risks and benefits associated with CCS, so the sample size was determined by that study's design.

sided alternative hypothesis with our study's average sample size per group and an alpha level of 0.05, I have 84% power to detect a small-to-medium ($d = 0.19$) effect.

Procedure

After providing informed consent, participants were asked about their concern about climate change⁷ and underwent an attention check. Participants were randomly assigned to the treatment ($n = 1,964$) or control ($n = 1,787$) arm of the survey. Treatment participants were shown one of the eight prompts shown in Table 2.1. Each prompt contained a question and a norm statement. The question was either positive or negatively framed (i.e., is this a commonly held norm? or: is this not a commonly held norm?) and each norm statement was either about corporate priorities or corporate practices and constructed as either an injunctive or descriptive statement. Control participants were randomly shown two of the four statements⁸ (topic x norm type combinations) and were asked their norm perception, with the question phrased in a neutral frame ("Do people agree or disagree with the statement below?")

After reading the prompt, participants were asked, "What is going through you mind as you consider your answer to the question whether most people [*agree or disagree*] with the statement? Please type each of your thoughts as it comes to mind into a box below. Put each thought into its own box." Participants were shown ten short text-response boxes and could type as many thoughts as they wanted. Each distinct thought was typed into a new box. Next, participants answered the norm perception question that reflects their condition, shown in Table 2.1.

⁷ Participants were asked, "How worried are you about global warming?" on a 4-point Likert scale of "Not at all worried", "Slightly worried", "Somewhat worried", and "Extremely worried".

⁸ Participants in the control condition were shown and response to two norm statements so that the length of the survey was closer to that of the treatment participants. Control participants did not do the thought listing or thought-categorizing tasks. Two of the four statements had significant order effects.

Participants then answer a series of risk and benefit perceptions about energy technology for the main survey, which are not relevant to this study. Participants were then shown each thought they listed during deliberation and rated the thought for whether it is, “evidence that most people would agree [that the statement is a norm] or that most people would disagree with the statement [that the statement is not a norm]?”. The response options were as follows: the thought is evidence that “the statement is a norm”, “the statement is not a norm”, “[the statement is] both”, or “[the statement is] neither”. Thoughts characterized as evidence “the statement is a norm” will be referred to as “norm-increasing” thoughts below, because they increase the perceived agreement with the norm; thoughts characterized as evidence “the statement is not a norm” are called “norm-decreasing”. Responses to this question allowed us to calculate two query-list metrics: query order and query-list content.

Measures

Social norm perception. The perceived social norm was measured as participants’ perception of how most people would agree or disagree with the statement (topic x norm type). The wording of the question posed agreed with the frame condition to emphasize the condition. In the neutral frame condition (shown to control participants), participants answered, “Do most people agree or disagree with the statement below?” on a 6-point Likert scale of “Almost everyone disagrees”, “Most people disagree”, “Some people disagree”, “Some people agree”, “Most people agree” and “Almost everyone agrees”. In the agree frame condition, participants answered, “Do you think the statement below is a belief commonly held by most people? Do most people agree with the statement?” on 6-point Likert scale “No, almost everyone disagrees”, “No, most people disagree”, “No, some people disagree”, “Yes, some people agree”, “Yes, most people agree”, and “Yes, almost everyone agrees”. In the disagree frame condition, participants answered, “Do you think

the statement below is a belief not commonly held by most people? Do most people disagree with the statement?” on a 6-point Likert scale of “Yes, almost everyone disagrees”, “Yes, most people disagree”, “Yes, some people disagree”, “No, some people agree”, “No, most people agree”, and “No, almost everyone agrees”. The response levels are the same in all three conditions, the only difference is the proceeding “Yes” or “No” to emphasize the frame condition. In all conditions, responses were coded so that agreement is positive, and disagreement is negative.

Query order. Query order was measured as the standardized median rank differences (SMRD) between norm-increasing and norm-decreasing arguments. SMRD is defined as $2(MR_i - MR_d)/n$, where MR_i is the median rank of arguments for norm-increasing in a respondent’s sequence of thought listing preceding the choice, MR_d is the median rank of arguments for norm-decreasing, and n is the total number of thoughts listed (Johnson et al., 2007). SMRD takes the value of +1 if all arguments for norm-increasing are recalled first, and -1 if all arguments for norm-decreasing are recalled first, and 0 if the two types of arguments do not cluster but are randomly interspersed.

Query-list content. Content is measured as the proportion of norm-increasing thoughts out of the total number of thoughts listed⁹.

Demographics. At the end of the survey, participants answered demographic questions about their age, income, level of education, gender, political orientation, political affiliation, and race. Participants in the location-targeted sample were also asked for their ZIP code. A detailed list of the demographic questions and response options is provided in Table A2.2.

Post-hoc response coding

⁹ I also calculated a version of query content that is the proportion of norm-increasing thoughts out of the number of norm-increasing and norm-decreasing thoughts, removing thoughts categorized as “both” or “neither” from the denominator of this metric. I did not include this metric in the main analysis because some participants do not list any norm-increasing or norm-decreasing thoughts and thus are dropped because their denominator is zero.

Post-hoc, two trained research assistants (RAs) categorized the aspects that treatment participants provided. Each aspect was coded for the following: (i) whether the memory is episodic [yes, no], (ii) whether the memory is semantic [yes, no], (iii) whether the aspect references a group or entity [yes, no], and what the references group or entity is [*text*], and (iv) whether the aspect references a social norm [yes, no]. All aspects that were longer than ten characters were read and categorized—ten characters was determined to be the minimum length of an aspect that could be interpreted after reviewing a random sample of 250 segments. Short aspects were too vague for researchers to interpret (i.e., “responsible”, “risks”, “concerned”). I addressed this in Study 2 by adding more classification questions for the participants so that all aspects could be coded by participants, without having to rely on interpretability or length of the aspect. There was a total of 6,122 thoughts listed by participants and 3,732 were categorized (3,622 were longer than ten characters and 110 of the short thoughts were categorized during the training process).

The RAs were first trained on the relevant psychological concepts, namely episodic and semantic memory and social norms, through assigned reading and discussion of the topics. Then, the two RAs categorized two sets of aspects ($n = 250$) and discussed all codes assigned to those aspects. The lead author and the two RAs developed and refined a coding manual with instructions, definitions, and examples for each category, see Table A2.3 for coding manual. Finally, I established interrater reliability. The RAs coded the same batch of aspects without consulting each other and their responses were compared. The first round ($n = 150$) established reliability on three of the four key categories—only the semantic memory coding failed. The lead author and RAs reviewed the full set together, amended and clarified the coding manual and ran a second round of

reliability testing ($n = 100$). All categories now reached sufficient interrater reliability (all $\kappa \geq 0.80$ ¹⁰), and the remainder of the aspects were coded. The coding manual is presented in Table A2.3.

Results

Norm Perceptions

In the control group (shown in green in Figure 2.1), the highest perceived norm endorsement was for the injunctive environmental practices norm ($M=4.49$, $SE=0.036$), followed by the descriptive economic priorities norm ($M=4.33$, $SE=0.040$), descriptive environmental practices norm ($M=3.87$, $SE=0.038$), and injunctive economic priorities norm ($M=3.04$, $SE=0.053$; Table A2.4). I used a mixed effects linear regression model to evaluate variation between norm perceptions, the model regressed norm perception on the norm (e.g., descriptive norm about economic priorities, injunctive norm about environmental practices), with random intercepts for participant ID. Since perceived norm endorsement is measured on a Likert-scale, I also ran the same model structure as a cumulative link mixed model. Perceived norm endorsement was significantly different between each of these four norms (Figure 2.1; Table 2.2).

For the treatment groups, the two environmental practice norms and the descriptive economic priorities norm showed a significant frame effect on norm perception—norm perceptions were significantly higher in the agree frame than in the disagree frame condition (p -values <0.0001 ; Table 2.3). Compared to control, the disagree frame significantly reduced norm perceptions (all p -values <0.001 ; Table 2.3). The agree frame was not significantly different from

¹⁰ The interrater reliability for each category is: episodic memory ($\kappa = 0.98$), semantic memory ($\kappa = 0.80$), contains referent ($\kappa = 0.88$), and contains social norm ($\kappa = 0.90$).

control for three of the statements (all p -values $> .10$; Table 2.3), suggesting that the natural frame by which people assess those norm statements is an agree frame.

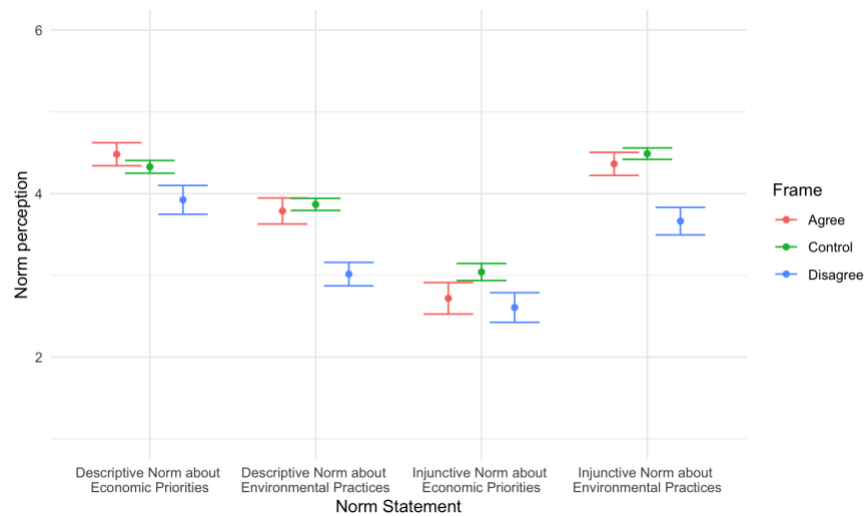


Figure 2.1. Average norm perception with 95% confidence interval. All differences are significant. Descriptive statistics for norm perceptions of the four statements (topic x norm type). The mean, standard error, and number of responses are shown for each frame (neutral control, agree frame, and disagree frame). 1= almost everyone disagrees, 3= some people disagree, 4=some people agree, 6 = almost everyone agrees.

<i>Predictors</i>	Norm Perception (numeric)			Norm Perception (ordinal)		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	4.32	4.24 – 4.41	<0.001			
Norm [Economic priorities, Injunctive]	-1.28	-1.40 – -1.17	<0.001	0.14	0.12 – 0.18	<0.001
Norm [Environmental practices, Descriptive]	-0.46	-0.57 – -0.34	<0.001	0.47	0.39 – 0.55	<0.001
Norm [Environmental practices, Injunctive]	0.17	0.06 – 0.28	0.003	1.27	1.06 – 1.51	0.008
Almost everyone disagrees!Most people disagree				0.02	0.02 – 0.03	<0.001
Most people disagree!Some people disagree				0.09	0.07 – 0.10	<0.001
Some people disagree!Some people agree				0.19	0.17 – 0.22	<0.001
Some people agree!Most people agree				1.15	1.01 – 1.32	0.035
Most people agree!Almost everyone agrees				6.27	5.33 – 7.38	<0.001
Random Effects						
σ^2	1.33			3.29		
τ_{00}	0.23 _{id}			0.48 _{id}		
ICC	0.15			0.13		
N	1787 _{id}			1787 _{id}		
Observations	3574			3574		
Marginal R ² / Conditional R ²	0.165 / 0.291			0.158 / 0.266		

Table 2.2. Two mixed effects regression models predicting norm perception by norm topic and type in the control group. The models predict norm perception as numeric values (left) and ordinal category (right). Norm perceptions is regressed on the norm (injunctive economic priorities, descriptive economic priorities, injunctive environmental priorities, and descriptive environmental priorities) with random effect for participant ID. The norm reference category is descriptive economic priorities.

Norm Topic	Norm Type	Neutral vs. Agree	Neutral vs. Disagree	Agree vs. Disagree
Economic priorities	Descriptive	t(413)=-1.88, p=.060	t(334)=4.10, p<.0001	t(456)=4.85, p<.0001
Economic priorities	Injunctive	t(413)=2.88, p=.004	t(428)=4.07, p<.0001	t(501)=0.83, p=.40
Environmental practices	Descriptive	t(356)=0.95, p=.37	t(389)=10.37, p<.0001	t(486)=7.07, p<.0001
Environmental practices	Injunctive	t(363)=1.56, p=.12	t(327)=8.86, p<.0001	t(464)=6.25, p<.0001

Table 2.3. Summary table showing the results of a series of t-tests between pairs of frames (neutral vs. agree, neutral vs. disagree, and agree vs. disagree). T-tests are performed individual for each norm topic and type pairing. The results of the t-test include the t-statistic, degrees of freedom, and the p-value.

Using linear regression and ordinal regression models, I investigated the effect of question frame (agree versus disagree) and norm type (descriptive versus injunctive) on norm perception (Table 2.4). The disagree frame significantly decreased perceived norm level compared to the agree frame for both norm topics ($\beta_{\text{economic priorities}} = -0.56, p < 0.001$; $OR_{\text{economic priorities}} = 0.53, p < 0.001$; $\beta_{\text{environmental practices}} = -0.77, p < 0.001$; $OR_{\text{environmental practices}} = 0.31, p < 0.001$). For the economic priorities topic, the injunctive formulation of the norm statement was lower than the perceived descriptive norm level ($\beta = -1.75, p < 0.001$; $OR = 0.11, p < 0.001$). The opposite pattern was observed for the environmental practices topic, for which the injunctive formulation was greater than the perceived descriptive norm level ($\beta = 0.57, p < 0.001$; $OR = 2.22, p < 0.001$). There was also a significant interaction between frame and norm type for the economic priorities topic: frame had a negative effect on descriptive norms but no significant effect on injunctive norms (Figure 2.2). I controlled for worry about climate change and found that, for both topics, participants who were more worried about climate change reported higher norm perceptions ($\beta_{\text{economic priorities}} = 0.09, p = 0.031$; $OR_{\text{economic priorities}} = 1.70, p = 0.019$; $\beta_{\text{environmental practices}} = 0.14, p < 0.001$; $OR_{\text{environmental practices}} = 1.24, p < 0.001$). Take together, participants perceive that companies fall short on their environmental practices (i.e., injunctive norm perceptions are higher than descriptive norm perceptions) and on economic priorities (i.e., descriptive norm perceptions are higher than injunctive norm perceptions).

<i>Predictors</i>	Economic Priorities Norm Perception (numeric)			Economic Priorities Norm Perception (ordinal)			Environmental Practices Norm Perception (numeric)			Environmental Practices Norm Perception (ordinal)		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
Intercept	4.23	3.95 – 4.52	<0.001				3.43	3.18 – 3.68	<0.001			
Frame [Disagree]	-0.56	-0.81 – -0.31	<0.001	0.53	0.39 – 0.72	<0.001	-0.77	-0.98 – -0.55	<0.001	0.31	0.23 – 0.43	<0.001
Norm type [Injunctive]	-1.75	-2.00 – -1.51	<0.001	0.11	0.08 – 0.16	<0.001	0.57	0.36 – 0.79	<0.001	2.22	1.61 – 3.06	<0.001
Worry abt climate change	0.09	0.01 – 0.18	0.031	1.13	1.01 – 1.26	0.028	0.14	0.06 – 0.21	<0.001	1.24	1.10 – 1.39	<0.001
Frame [Disagree] x Norm type [Injunctive]	0.44	0.10 – 0.79	0.012	1.70	1.09 – 2.65	0.019	0.08	-0.23 – 0.39	0.607	1.22	0.78 – 1.92	0.386
Almost everyone disagrees/Some people disagree				0.05	0.03 – 0.08	<0.001				0.06	0.04 – 0.10	<0.001
Most people disagree/Some people agree				0.18	0.12 – 0.26	<0.001				0.31	0.21 – 0.46	<0.001
Some people disagree/Some people agree				0.37	0.26 – 0.54	<0.001				1.07	0.73 – 1.55	0.733
Some people agree/Some people agree				1.12	0.78 – 1.61	0.551				4.27	2.90 – 6.30	<0.001
Most people agree/Almost everyone agrees				5.11	3.44 – 7.58	<0.001				20.44	13.20 – 31.65	<0.001
Observations	990			990			974			974		
R ² / R ² adjusted	0.251 / 0.247			0.247			0.146 / 0.142			0.153		

Table 2.4. Four regression models predicting norm perception. Regressions are run for the two norm topics separately (economic priorities and environmental practices). For each topic, a linear regression on the numeric variable and an ordinal regression on the ordinal variable are run. Norm perception is regressed on frame, norm type, the interaction between frame and norm type, and worry about climate change. The reference category for frame is the agree frame and the reference category for the norm type is descriptive. From left to right, the models are: linear regression predicting economic priorities norm, ordinal regression predicting economic priorities norm, linear regression predicting environmental practices norm, and ordinal regression predicting environmental practices norm.

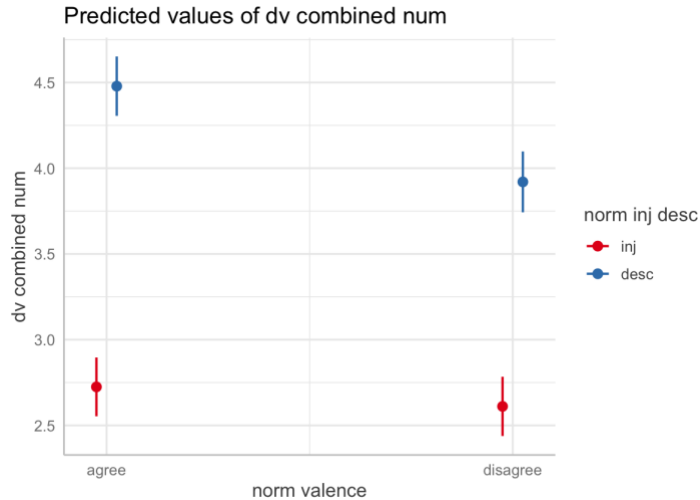


Figure 2.2. Interaction plot for the linear regression model predicting economic priorities norm, Table 2.4. The plot shows how the relationship between frame (x-axis) and norm perception (y-axis) varies by norm type (color; red = injunctive norm, blue = descriptive norm).

Query metrics

Next, I examined the cognitive mechanism underlying norm estimation for the agree versus disagree frame and then for the injunctive versus descriptive norm construction. Query order is measured as SMRD and query content is measured as proportion of norm-increasing thoughts¹¹. Norm-increasing thoughts were queried and thus listed earlier than norm-decreasing thoughts—as indicated by positive SMRDs in the descriptive economic priorities norm, the descriptive environmental practices norm, and the injunctive environmental practices norm. Participants began their deliberation by considering why these three topics might be normative. The injunctive economic priorities norm had a negative SMRD, indicating norm-decreasing were valued first. These patterns are consistent between the agree and disagree frames (Table 2.5).

¹¹ SMRD and proportion of norm-increasing thoughts are positively correlated (69%) in this study, as predicted by query theory.

The mean proportion of norm-increasing thoughts out of all aspects listed is 0.38 ($SD = 0.41$) and the mean proportion of “norm-increasing” thoughts out of norm-related aspects (i.e., aspects that are a reason why the statement is norm-increasing or decreasing) is 0.64. ($SD = 0.43$). In all conditions, there are more norm-increasing thoughts than norm-decreasing thoughts: the injunctive economic priorities statement with the disagree frame has the lower proportion ($M = 0.52$, $SD = 0.44$).

Topic	Norm Type	Agree frame				Disagree frame			
		Order (SMRD)		Prop. norm-increasing thoughts		Order (SMRD)		Prop. norm-increasing thoughts	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Economic priorities	Descriptive	0.361	0.0035	0.445	0.0026	0.329	0.0037	0.455	0.0028
Economic priorities	Injunctive	-0.003	0.0036	0.325	0.0025	-0.014	0.0037	0.316	0.0025
Environmental practices	Descriptive	0.184	0.0037	0.378	0.0026	0.097	0.0036	0.331	0.0025
Environmental practices	Injunctive	0.230	0.0037	0.406	0.0027	0.215	0.0037	0.382	0.0027

Table 2.5. Summary table of the query metrics, order and proportion, for norm topic, type, and frame. Query order is measured as SMRD, standardized median rank difference, where -1 represents all norm-decreasing thoughts clustering first, 0 represents random dispersion, and 1 represents all norm-increasing thoughts clustering first. Proportion is calculated as the number of norm-increasing thoughts divided by the total number of thoughts listed. The mean and standard error is reported for each metric and category.

To examine the conditions that affect query order and content, I regressed SMRD and proportion of norm-increasing thoughts on frame, norm type, topic, and the two-way interactions (Table 2.6). Frame did not affect query order nor proportion ($p\text{-values} > 0.05$). There was a significant interaction between norm type and topic ($\beta_{\text{SMRD}} = 0.44$, $t = 6.05$, $p < 0.001$; $\beta_{\text{SMRD}} = 0.44$, $t = 4.56$, $p < 0.001$). The interactions suggest that query metrics vary by norm type for the

economic priorities norm and not for the environmental practices norm (Figure 2.3). Subsequent analysis on the economic priorities norm, finds that SMRD and proportion are both significantly lower for the injunctive norm type than for the descriptive norm type (SMRD: $t(985) = -7.05, p < 0.001$; Prop.: $t(979) = -4.93, p < 0.001$).

<i>Predictors</i>	SMRD			Prop. of norm-inc. thoughts		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.37	0.27 – 0.46	<0.001	0.45	0.40 – 0.50	<0.001
Frame: Disagree	-0.04	-0.17 – 0.08	0.474	-0.00	-0.06 – 0.06	0.979
Type: Injunctive	-0.38	-0.50 – -0.26	<0.001	-0.13	-0.19 – -0.07	<0.001
Topic: Envi. Prac.	-0.19	-0.31 – -0.07	0.002	-0.08	-0.14 – -0.01	0.016
Frame: Disagree x Type: Injunctive	0.05	-0.10 – 0.19	0.523	0.00	-0.07 – 0.07	0.948
Type: Injunctive x Topic: Envi. Prac.	0.44	0.29 – 0.58	<0.001	0.17	0.10 – 0.24	<0.001
Frame: Disagree x Topic: Envi. Prac.	-0.03	-0.17 – 0.11	0.679	-0.04	-0.11 – 0.04	0.334
Observations	1964			1964		
R ² / R ² adjusted	0.026 / 0.024			0.015 / 0.012		

Table 2.6. Two linear regressions predicting SMRD (left) and proportion of norm-increasing thoughts (right). Query metrics are regressed on frame, norm type, topic, and the two-way interactions. The reference category for frame is the agree frame, the reference category for the norm type is descriptive, and the reference category for topic is economic priorities.

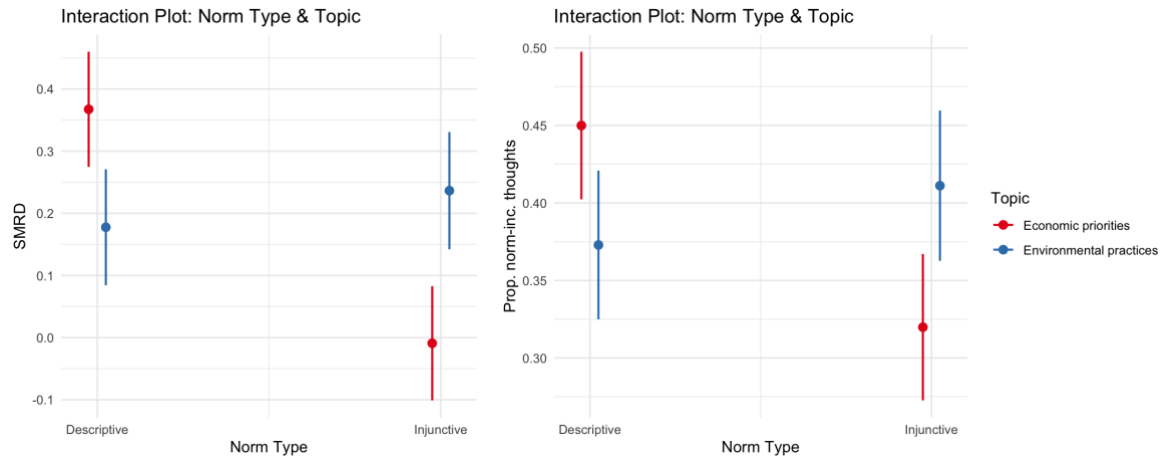


Figure 2.3. Two interaction plots for the linear regression models predicting SMRD (left) and proportion of norm-increasing thoughts (right), Table 2.6. The plots show how the relationship between norm type (x-axis) and the respective query metrics (y-axis) varies by topic (color; red = economic priorities, blue = environmental practices).

Next, I examined whether the query metrics predicted norm perception. Using linear regression, I regressed norm perception on SMRD, proportion of norm-increasing thoughts, worry about climate change, frame, norm type, and the interaction between frame and norm type. Topics were evaluated in separate models. Query order (SMRD) was a significant predictor of norm perceptions in both models and query content (proportion of thoughts) was not. Higher SMRD, which indicates norm-supporting thoughts clustered earlier than norm-opposing thoughts, predicted higher perceived norms for economic priorities ($b = 0.18, p = 0.018$) and environmental practices ($b = 0.26, p < 0.001$). In the present study, query order was more predictive of norm perceptions even though both order and content were affected by the stimuli.

<i>Predictors</i>	Economic Priorities Norm Perception			Environmental Practices Norm Perception		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
Intercept	4.21	3.91 – 4.51	<0.001	3.39	3.13 – 3.64	<0.001
SMRD	0.18	0.03 – 0.33	0.018	0.26	0.13 – 0.39	<0.001
Prop. norm-supporting thoughts	-0.06	-0.35 – 0.23	0.678	-0.04	-0.29 – 0.22	0.781
Frame [Disagree]	-0.55	-0.80 – -0.30	<0.001	-0.75	-0.96 – -0.54	<0.001
Norm type [Injunctive]	-1.70	-1.94 – -1.45	<0.001	0.56	0.35 – 0.78	<0.001
Worry abt climate change	0.09	0.00 – 0.17	0.041	0.14	0.06 – 0.21	<0.001
Frame [Disagree] x Norm type [Injunctive]	0.44	0.09 – 0.79	0.013	0.06	-0.24 – 0.36	0.686
Observations	990			974		
R ² / R ² adjusted	0.257 / 0.252			0.169 / 0.164		

Table 2.7. Two linear regressions predicting norm perception. Regressions are run for the two norm topics separately (economic priorities and environmental practices). For each topic, norm perception is regressed on SMRD, proportion of norm-increasing thoughts, frame, norm type, the interaction between frame and norm type, and worry about climate change. The reference category for frame is the agree frame and the reference category for the norm type is descriptive.

To investigate the query theory mechanism, I ran a mediation analysis on the economic priority norm condition. This was the only norm topic for which a condition (norm type) affected the query metrics. For the economic priority norm conditions, mediation analyses were conducted to examine whether query order or content mediated the relationship between norm type (injunctive versus descriptive) and norm perception. I controlled for frame in all analyses. The analyses were conducted using the mediation package in R, with 10,000 bootstrap resamples to compute confidence intervals for the indirect effect (Tingley et al., 2014). The total effect of norm type on norm perception was significant, with an unstandardized coefficient of $\beta = -1.54$ (95% CI[-1.73,-1.37], $p < 0.0001$). The direct effect of was also significant, $\beta = -1.48$ (95% CI [-1.67, -1.30], $p < 0.00001$). The indirect effect, representing the effect of norm type on norm perception through order, was significant, with ACME = -0.058 (95% CI [-0.10,-0.02], $p = 0.004$). This

suggests that order partially mediated the effect of norm type on norm perception for norms about economic priorities. For proportion of norm-increasing thoughts, the indirect effect was not significant, with ACME = -0.024 (95% CI [-0.06,0.00], $p = 0.098$). This suggests that proportion did not mediate the effect of norm type on norm perception.

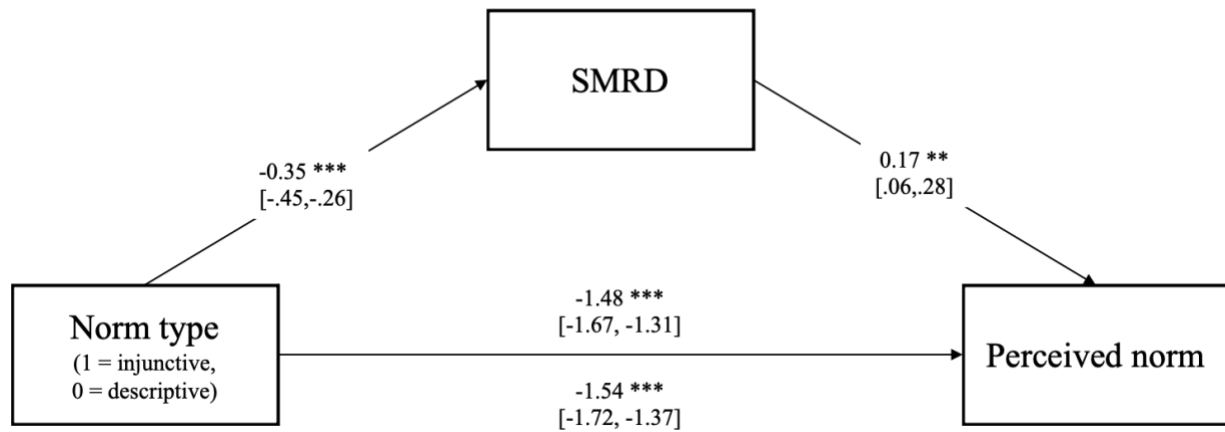


Figure 2.4. Mediation analysis testing indirect effect of SMRD on norm perception, controlling for frame. The 95% confidence intervals are calculated using 10,000 bootstrap resamples.

Memory features

Across the 3,732 thoughts that were categorized, 634 referenced a memory. Of the aspects that likely reference a memory, there were 610 that contained a semantic memory and only 17 thoughts that contained an episodic memory. Semantic memories most often referenced companies ($n = 240$) and people ($n = 110$). For example, an aspect that was categorized as a semantic memory is “most people are concerned about pollution” because the statement references a pattern of observed behavior. “I read the earth’s temperature went up 1.5 degrees last year” was categorized as an episodic memory because the participant referenced a specific event in the past, reading about global temperatures. Coding the aspects required Additional examples are shown in Table 2.8.

Thought Aspect	Length	Memory	Episodic Mem.	Semantic Mem.	Ref. Grp.	Ref. Grp. Detail	Social Norm
More people need knowledge about this	37	3	0	1	1	people	1
I think the younger generation would like to care more about the planet.	72	3	0	1	1	younger generation	1
All companies prioritize their bottom line in America	53	3	0	1	1	companies	1
Most of the snow here has melted.	33	3	1	0	0	NA	0
waste of time	13	1	0	0	0	NA	0
Evil	4	NA	NA	NA	NA	NA	NA
I read the earth's temperature went up 1.5 degrees last year.	61	3	1	0	0	NA	0
companies still need to make a profit to stay in business	57	2	0	1	1	companies	0
most people are concerned about pollution	41	3	0	1	1	people	1
It's a huge issue that most don't pay attention to	48	3	0	1	1	[people]	1
Petroleum products are everywhere	33	3	0	1	0	NA	0
global warming is bad	21	1	0	0	0	NA	0
Some people can't afford being environmentally friendly.	56	3	0	1	1	people	0

Table 2.8. Example thoughts from Study 1 thought lists with ad hoc coding. Each thought is evaluated for length (number of characters), whether the thought is a memory (1=not a memory, 2=unclear, 3=is a memory), if the thought contains an episodic memory (1=yes, 0=no), if the thought contains a semantic memory (1=yes, 0=no), if there is a reference group mentioned (1=yes, 0=no), what the reference group is, and if there is a social norm mentioned (1=yes, 0=no).

Social norms appeared in 488 of the categorized thoughts. For example, “It’s a huge issue that most don’t pay attention to” was categorized as a social norm perception because the participant thinks there is a descriptive norm of not paying attention to corporate economic priorities. Another example is, “I think the younger generation would like to care more about the planet” because it describes a norm about the younger generation (Table 2.8). A Wilcoxon rank

sum test revealed a significantly higher social norm prominence in descriptive norm conditions than in injunctive norm conditions, with a test statistic of $W = 58236, p < 0.001$.

Discussion

The results from Study 1 show norm perceptions varied significantly across the four norms (topic x norm type) and the agree (versus disagree) frame significantly increased perceptions. Counter to our intuition¹², frame did not affect query order or content. Norm construction affected the query metrics only in the economic priorities domain. An exploratory mediation analysis on the economic priorities norms found that query order partially mediated the relationship between norm type and norm perception. The query order for the injunctive economic priorities norm was the only norm of the four where the aspects against the norm clustered earlier than the aspects for the norm. These results may be driven by the social versus antisocial connotations of the two norm topics: environmental practice was pro-social (e.g., “Energy companies should adopt practices to reduce the environmental impact of their operations, even when it is costly or risky to them”), while economic priorities were anti-social (e.g., “Energy companies should prioritize their own bottom line, even if it increases (worsens) health, environmental, and social risks for the community”). The injunctive versus descriptive focus for the anti-social norm—whether companies prioritize their bottom line and if companies *should* prioritize their bottom line—had a significant effect on query order. In the anti-social injunctive norm condition, the first aspects were about why companies should *not* prioritize their bottom line. This suggests a strong aversion to this principle and that the natural query order for injunctive norms is different for social versus anti-social norms, such that the pro-social option is attended to first. This does not seem to be the

¹² There were no formal hypotheses for Study 1.

case with descriptive norms, where norm-increasing reasons are evaluated first in both social and antisocial norm conditions.

Study 2: Social norms about choosing meat-based vs. vegetarian meals

Method

To better measure these effects, I selected a second norm topic—a choice between a meat-based or vegetarian meal—to provide participants with a more familiar topic likely to invoke more frequent and recent personal experiences to draw on. Following the design of Study 1, recruited participants listed their thoughts as they deliberated what was normative in a common choice scenario before making a final determination and categorizing each thought. Participants were asked to think about the choice between a meat-based meal and a vegetarian meal, in a 2x2 between-subjects design where norm type (injunctive versus descriptive) and frame (meat-based vs. vegetarian) were randomly assigned. The full stimuli text is shown in Table 2.9. Study 2 was preregistered (As.Predicted #207418).

Participants

A power analysis indicated that a sample of 540 was needed to detect a small-to-medium effect ($f = .15$, determined by Study 1) of frame on norm perception with 80% power at the .05 level. I recruited 579 participants to this study through Prolific Academic. The sample, collected in January 2025, was demographically representative of the United States based on age, sex, and political affiliation. As preregistered, participants were excluded from analysis if they failed an attention check ($n = 24$) or spent longer than 60 minutes on the survey ($n = 15$). After removing ineligible participants, I was left with 540 participants. 50% of all samples identified as female and the median age group was 35-44 years old. The median annual household income bracket is

\$50,000-74,999, and 50% of respondents had received a college degree. The distribution of political affiliation is 31% Democrat, 36% Independent, 30% Republican, and 3% other. Full demographic information is presented in Tables A2.6.

Procedure

After participants provided informed consent and passed an attention check question, they were given a practice task to learn the survey mechanism for thought listing (shown in Figure A2.1). Next, all participants read an introduction, “Now, we would like you to think about a common decision—what to eat for lunch. We want to know how most people would make this decision. Imagine a person, somebody like you, is at a fast casual restaurant for lunch with some new friends and the restaurant offers two meal choices—a meat meal and a vegetarian meal. The meals have comparable side dishes and flavor profiles, the only difference is that the protein in the meat meal is chicken and the protein in the vegetarian meal is beans and seitan (a soy-based protein).” Participants were randomly assigned to one of four prompts based on their condition, the full text for each is shown in Table 2.9. To direct attention to the appropriate descriptive or injunctive norm, we ask about dinner behavior (descriptive norm) and friend opinion (injunctive norm). Friends were used for the injunctive norm to indicate the diner would be eating with those people. Then participants did the same thought listing task used in Study 1. After listing all the thoughts that went through their mind, participants answered their prompt question (Table 2.9). Next, participants indicated what they thought the most common reasons were that people would eat less meat.

Frame	Norm Type	Stimuli Text
Meat meal	Descriptive	Out of 10 diners making this choice, how many diners would pick the meat-based meal?

Vegetarian meal	Descriptive	Out of 10 diners making this choice, how many diners would pick the vegetarian meal?
Meat meal	Injunctive	Out of 10 friends, how many friends expect the diner to choose the meat-based meal?
Vegetarian meal	Injunctive	Out of 10 friends, how many friends expect the diner to choose the vegetarian meal?

Table 2.9. Detailed stimuli for Study 2.

After answering the dependent variable questions, participants were shown each thought they had listed in the order in which they had entered them. For each, participants answered, “did this thought increase or decrease your estimate?” with options “it increased my estimate” (termed estimate-increasing), “it decreased my estimate” (termed estimate-decreasing), “it did both”, and “it did neither”. For each thought, participants also categorized the type of information, “what type of information is in this thought?”, by selecting all that apply from options “this is a memory or reference to past events”, “this is an opinion”, “this is a factor or piece of information”, “this is a question”, and “other”. For thoughts that were categorized as a memory, an opinion, or piece of information, participants were asked “is this thought based on a specific or multiple memories” with options for “specific memory”, “multiple memories/a pattern”, and “neither”. After all thoughts were categorized, participants answered a second pair of attention checks and the same demographic questions used in Study 1.

Measures

Social norm perception. Norm perceptions were measured in response to the four condition-based questions presented in Table 2.9. All estimates were on a scale from 0 to 10 where 0 represents no one and 10 represents everyone. Responses were recoded such that in all conditions higher responses indicated more vegetarian meals.

Motivation for eating less meat. Motivation for eating less meat was assessed with two questions. In the first question, participants selected all reasons people eat less meat from the list: “health”,

“cost”, “environmental impact”, “flavor/taste”, “ease of cooking”, “religious or moral reasons”, and “other (please describe)”. In the second question, participants selected the “one, most common, reason people eat less meat” from the same list.

Query order. Query order was measured as the standardized median rank differences (SMRD) between vegetarian meal and meat meal arguments. In this study, thoughts were classified as ‘vegetarian-meal arguments’ if they are in favor of picking the vegetarian meal or against picking the meat meal. ‘Meat-meal arguments’ were those arguments in favor of picking the meat meal or against picking the vegetarian meal. SMRD takes the value of +1 if all arguments for vegetarian meals are recalled first, and -1 if all arguments for meat meals are recalled first, and 0 if the two types of arguments are randomly interspersed.

Query-list content. Content is calculated for each participant as the proportion of vegetarian-meal arguments out of the total number of thoughts listed.

Semantic memory content. Semantic memory content is calculated for each participant as the proportion of semantic memories out of the total number of thoughts listed.

Episodic memory content. Episodic memory content is calculated for each participant as the proportion of episodic memories out of the total number of thoughts listed.

Demographics. Participants answered the same demographic questions used in Study 1. A detailed list of the demographic questions and response options is provided in Table A2.2.

Results

Norm perceptions

Across framing conditions, the average perceived descriptive norm was that 2.96 diners out of 10 would pick the vegetarian meal ($SD = 1.93$) and the average perceived injunctive norm

was that 2.64 friends out of 10 think the diner should pick the vegetarian meal ($SD = 2.21$). These two norm prevalence perceptions are not significantly different ($F(1,538) = 3.21, p = 0.074$). The magnitudes of the perceived norms by frame (meat vs. vegetarian) are shown in Table 2.10. The descriptive norm was higher in the vegetarian frame than in the meat frame, but the difference was not significant ($p = 0.22$). The injunctive norm was higher in the meat frame, but the difference was also not significant ($p = 0.25$). In the vegetarian frame, the descriptive norm was significantly higher than the injunctive norm ($F(1,271) = 6.80, p = 0.0096$). In the meat frame, the injunctive and descriptive norm were not significantly different ($p = 0.94$). This suggests there was agreement across the two frames about the descriptive norm and across norm type in the meat frame. In the vegetarian frame, the descriptive norm being greater than the injunctive norm suggests respondents think that people eat more vegetarian meals than their peers/friends would expect.

Frame	Norm Type	Norm Perception	
		<i>M</i>	<i>SE</i>
Meat	Descriptive	2.81	0.17
Meat	Injunctive	2.80	0.21
Vegetarian	Descriptive	3.09	0.15
Vegetarian	Injunctive	2.48	0.18

Table 2.10. Average norm perceptions by frame and norm type. Mean and standard error are reported. Responses are on a scale from 0 to 10 where 0 represents no people endorse the vegetarian meal and 10 represents all people endorse the vegetarian meal.

As shown in Table 2.11, participants reported that health was the most common reason people eat less meat ($n = 255$), followed by religious or moral reasons ($n = 191$), cost ($n = 46$), environmental impact ($n = 33$), and taste ($n = 6$). The order of reasons was slightly different for the multi-select reason question. Health and religion were still most selected, but environmental impact ($n = 331$) was selected more than cost ($n = 213$), suggesting environmental impact is more prominent in a multi-motivation framework but cost is more important than environmental impact

in a single-motivation framework. The prominence of different reasons was not affected by condition, as determined by a series of two-way (norm type x frame) ANOVA tests (*all p-values* > 0.10).

Reasons people eat less meat		
<i>Reason</i>	<i>"most common" (single choice)</i>	<i>"a reason" (multi-choice)</i>
Health	255	468
Religions or moral reasons	191	442
Cost	46	213
Environmental impact	33	331
Flavor/taste	6	108
Other	9	28
Ease of cooking	0	67

Table 2.11. Summary of why people eat less meat. Results are reported as the number of participants ($n = 540$) who selected each reason.

Query theory metrics

To examine the cognitive process of norm estimation, I measured condition effects on the query theory mechanisms—order and proportion of thoughts. In this study, query order was measured as SMRD, and higher values indicate vegetarian-meal thoughts (either pro-vegetarian or anti-meat, depending on the frame) occurred earlier in the aspect list than meat thoughts. In both meat frame conditions and the descriptive vegetarian condition, SMRD was negative, meaning meat thoughts clustered earlier than vegetarian thoughts. The significant clustering of thoughts in the meat frames is validation of query theory’s memory-retrieval mechanisms. In these conditions, the first query elicited thoughts either about why someone would pick the meat meal (in the meat condition) or why someone would not pick the vegetarian meal (in the vegetarian condition) (Table 2.12). The positive SMRD in the injunctive vegetarian norm condition revealed that the thoughts elicited by the first query tended to focus on why someone would pick the vegetarian meal. I used

linear regression to evaluate the effect of frame and norm type on SMRD. I regressed SMRD on frame, norm type, and the interaction between frame and norm type (Table 2.13). There was a significant main effect of frame on SMRD ($\beta = 0.32, p < 0.0001$), such that SMRD is higher in vegetarian frame than in the meat frame. There was not a significant effect of norm type on SMRD, indicating that the injunctive (versus descriptive) construction of the norm had the same query order (Table 2.13a).

Across all conditions, the average proportion of vegetarian thoughts was 0.22 (SD = 0.31) and the average proportion of meat thoughts was 0.38 (SD = 0.41). Proportion of vegetarian meal thoughts was higher in the vegetarian frame than in the meat frame ($t(491)=5.08, p < 0.001$). And the proportion of meat-meal thoughts was higher in the meat frame than in the vegetarian frame ($t(531)=4.63, p < 0.001$). Neither proportion metric varied significantly by norm type (prop. veg thoughts: $t(517)=0.19, p = 0.85$; prop. meat thoughts: $t(530)=0.26, p = 0.80$). To investigate the effect of frame and norm type on the two proportion metrics, each metric was regressed on frame and norm type. As with SMRD, proportion of vegetarian thoughts is predicted by frame, such that participants in the vegetarian frame have higher proportions than those in the meat frame ($\beta = 0.13, p < 0.001$; Table 2.13b). In the proportion of meat thoughts model, there is a significant interaction between frame and norm type ($\beta = -0.14, p = 0.048$; Table 2.13c). The effect of vegetarian frame, compared to meat frame, had a negative impact on both types of norms but the effect is more negative for the injunctive norm (Figure 2.5). SMRD is 50% correlated with proportion of pro-vegetarian thoughts and 55% correlated with proportion of pro-meat thoughts across all four norms.

Frame	Norm Type	SMRD		Prop. pro-veg. thoughts		Prop. pro-meat thoughts	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Meat	Descriptive	-0.37	0.059	0.15	0.021	0.42	0.035

Meat	Injunctive	-0.37	0.062	0.15	0.022	0.50	0.038
Vegetarian	Descriptive	-0.051	0.067	0.27	0.027	0.33	0.033
Vegetarian	Injunctive	0.065	0.071	0.29	0.034	0.26	0.033

Table 2.12. Summary table of query metrics by norm—SMRD, proportion of vegetarian-increasing thoughts, and proportion of meat-increasing thoughts. SMRD, standardized median rank difference, ranges from -1 to 1 where -1 represents all meat-increasing thoughts clustering first, 0 represents random dispersion, and 1 represents all vegetarian-increasing thoughts clustering first. Proportion is calculated as the number of vegetarian-increasing or meat-increasing thoughts divided by the total number of thoughts listed. The mean and standard error is reported for each metric and category.

<i>Predictors</i>	SMRD			Prop. of veg-meal thoughts			Prop. of meat-meal thoughts		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	-0.37	-0.50 – -0.24	<0.001	0.15	0.09 – 0.20	<0.001	0.42	0.36 – 0.49	<0.001
Frame: Vegetarian	0.32	0.14 – 0.50	<0.001	0.13	0.06 – 0.20	<0.001	-0.10	-0.19 – -0.00	0.047
Type: Injunctive	-0.00	-0.18 – 0.18	0.982	0.00	-0.07 – 0.08	0.914	0.07	-0.02 – 0.17	0.142
Frame: Vegetarian x Type: Injunctive	0.12	-0.14 – 0.37	0.363	0.01	-0.09 – 0.12	0.819	-0.14	-0.27 – -0.00	0.048
Observations	540			540			540		
R ² / R ² adjusted	0.061 / 0.056			0.046 / 0.040			0.045 / 0.040		

Table 2.13. Three linear regression models predicting query metrics—SMRD (left, a), proportion of vegetarian-increasing thoughts (center, b), and proportion of meat-increasing thoughts (right, c). Query metrics are regressed on frame, norm type, and the interaction between frame and norm type. The reference category for frame is the meat frame and the reference category for norm type is the descriptive norm construction.

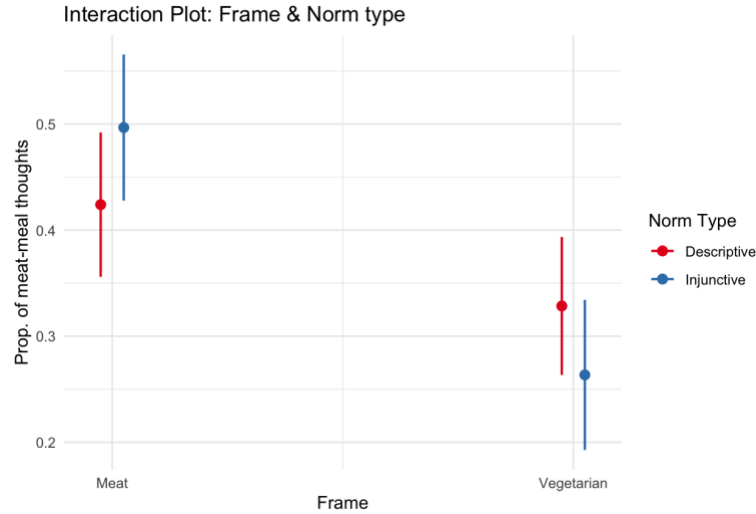


Figure 2.5. Interaction plot for the linear regression model predicting proportion of meat-increasing thoughts, Table 2.13. The plot shows how the relationship between frame (x-axis) and thought proportion (y-axis) varies by norm type (color; red = descriptive norm, blue = injunctive norm).

Episodic & semantic memory

Next, I investigated the memory features of the aspect list, specifically the proportion of semantic and episodic memories. Overall, semantic memories were more prominent than episodic memories ($M_{\text{semantic}} = 0.67$, $SD_{\text{semantic}} = 0.41$, $M_{\text{episodic}} = 0.16$, $SD_{\text{episodic}} = 0.30$; $t(985) = 24.41$, $p < 0.001$). Table 2.14 shows the prominence of semantic and episodic memories by condition. I investigated the effect of frame and norm type on proportion of semantic and episodic memories using linear regression models, shown in Table 2.15. There was not a significant main effect of frame or norm type on either type of memory. There was a significant interaction between frame and norm type predicting proportion of semantic memories (Figure 2.6). The interaction can be interpreted to mean that the frame (vegetarian compared to meat) affects the proportion of semantic memories listed positively for descriptive norms and negatively for injunctive norms. The effect

of the vegetarian frame (compared to the meat frame) on injunctive norm perception was directionally significant ($\beta = -0.089$, $p=0.094$) and the effect of vegetarian frame on descriptive norm perception was not significant ($\beta = 0.054$, $p = 0.25$).

Frame	Norm Type	Prop. semantic memories		Prop. episodic memories	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Meat	Descriptive	0.65	0.033	0.16	0.026
Meat	Injunctive	0.75	0.039	0.16	0.022
Vegetarian	Descriptive	0.70	0.032	0.15	0.023
Vegetarian	Injunctive	0.66	0.036	0.17	0.031

Table 2.14. Summary table of proportion of semantic and episodic memories by norm frame and norm type. Mean and standard error are reported, proportion ranges from 0 to 1.

Predictors	Prop. Semantic Memory			Prop. Episodic Memory		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.65	0.58 – 0.72	<0.001	0.16	0.11 – 0.21	<0.001
Frame [Vegetarian]	0.05	-0.04 – 0.15	0.268	-0.01	-0.08 – 0.06	0.793
Norm type [Injunctive]	0.10	-0.00 – 0.19	0.053	0.00	-0.07 – 0.07	0.955
Frame [Vegetarian] x Norm type [Injunctive]	-0.14	-0.28 – -0.00	0.042	0.02	-0.08 – 0.12	0.663
Observations	540			540		
R ² / R ² adjusted	0.009 / 0.003			0.001 / -0.005		

Table 2.15. Two linear regressions predicting the proportion of semantic memories (left) and episodic memories (right). Proportion is regressed on frame, norm type, and the interaction between frame and norm type. The reference category for frame is the meat frame and the reference category for norm type is the descriptive norm construction.

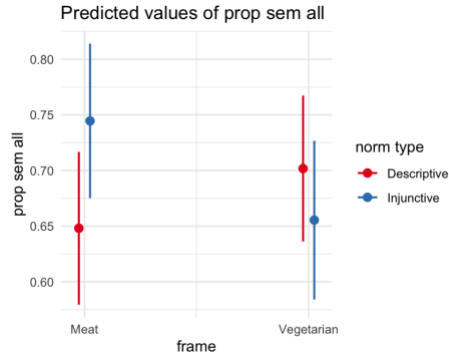


Figure 2.6. Interaction plot for the linear regression model predicting proportion of semantic memories, Table 2.15. The plot shows how the relationship between frame (x-axis) and memory proportion (y-axis) varies by norm type (color; red = descriptive norm, blue = injunctive norm).

Mechanisms of norm construction: query theory metrics & memory type

Finally, I investigated whether any of the cognitive mechanisms predicted norm perception, using linear regression analysis. In separate regressions for each norm, I regressed norm perception on SMRD, proportion of pro-vegetarian thoughts, proportion of pro-meat thoughts, and proportion of semantic memories (Table 2.16). The mechanisms that significantly predicted norm perceptions varied by norm. The descriptive meat norm was predicted by SMRD ($\beta = 0.67, p = 0.039$) and proportion of vegetarian-meal thoughts ($\beta = 2.14, p = 0.004$); higher SMRD and higher proportion were both associated with higher estimates of vegetarian meal choice in this norm condition. The injunctive meat norm was predicted by proportion of meat-meal thoughts ($\beta = -1.58, p = 0.004$) and proportion of semantic memories ($\beta = -1.06, p = 0.018$). Both estimates are negative, suggesting that a higher proportion of meat-meal thoughts and semantic memories were associated with seeing the vegetarian meal as less normative. Both the descriptive and injunctive vegetarian norms were predicted by proportion of vegetarian-meal thoughts ($\beta_{\text{desc. veg. norm}} = 1.10, p_{\text{desc. veg. norm}} = 0.029$; $\beta_{\text{inj. veg. norm}} = 1.76, p_{\text{inj. veg. norm}} = 0.0002$), such that higher proportion of thoughts was associated with higher norm estimates.

<i>Predictors</i>	Desc. Meat Norm Percep.			Inj. Meat Norm Percep.			Desc. Vegetarian Norm Percep.			Inj. Vegetarian Norm Percep.		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	2.52	1.84 – 3.21	<0.001	4.52	3.58 – 5.46	<0.001	2.91	2.22 – 3.61	<0.001	2.23	1.53 – 2.92	<0.001
SMRD	0.67	0.03 – 1.31	0.039	0.10	-0.63 – 0.83	0.783	0.17	-0.27 – 0.60	0.443	0.45	-0.19 – 1.08	0.168
Prop. veg thoughts	2.14	0.70 – 3.57	0.004	-0.72	-2.47 – 1.04	0.421	1.10	0.11 – 2.08	0.029	1.76	0.64 – 2.89	0.002
Prop. meat thoughts	0.45	-0.57 – 1.48	0.385	-1.58	-2.65 – -0.52	0.004	-0.62	-1.48 – 0.25	0.159	-0.04	-1.13 – 1.05	0.939
Prop. semantic memories	0.06	-0.81 – 0.92	0.899	-1.06	-1.94 – -0.19	0.018	0.13	-0.65 – 0.92	0.738	-0.42	-1.23 – 0.39	0.305
Observations	135			132			148			125		
R ² / R ² adjusted	0.158 / 0.132			0.130 / 0.102			0.086 / 0.060			0.224 / 0.198		

Table 2.16. Four linear regressions predicting norm perception (descriptive meat norm, injunctive meat norm, descriptive vegetarian norm, and injunctive vegetarian norm). Norm perception is regressed on SMRD, proportion of vegetarian-increasing thoughts, proportion of meat-increasing thoughts, and proportion of semantic memories.

Discussion

In Study 2, I did not find a significant frame effect. This may be due to the topic of meal choice: though this topic was selected as potentially giving participants an abundance of personal experiences to draw from, participants may also have perceived the decision as low-impact and thus searched their memory for fewer exemplars. The significant effect of norm type within the vegetarian frame may be driven by the health benefits of a plant-based diet, which is a more individual than collective motivation. The data showed that the perceived descriptive norm is higher than the perceived injunctive norm, such that people think it is more common to eat a vegetarian meal than to think everyone should eat vegetarian meals. Interestingly, the query order analysis revealed that the vegetarian injunctive norm condition was the only group that began by thinking about why to choose the vegetarian meal. This suggests that meat consumption is the implicit default—all other participants, even in the vegetarian frame, started by considering the

meat meal. There was a significant effect of frame on query order and content, in support for *H1a* and *H1b*. Norm type did not affect query order or content, so I reject *H2a* and *H2b*. Although there was not a main effect of frame on norm perception, query order and content did predict norm perceptions. The mechanism varied by norm topic, *H3a* was only supported in the descriptive meat norm condition and *H3b* was supported in all conditions. Finally, in the exploratory memory analysis, I found that semantic memories were more prominent in all conditions. The memory type analysis suggests that semantic memories are more dominant in the social cognition of norms

General Discussion

This chapter found that query order and content were significant predictors of norm perceptions, although which metric was more important varied by norm topic. This suggests there is some ad hoc construction or memory-retrieval processing that is related to norm perceptions, supporting the basic premises of norm theory and query theory in this context (Kahneman & Miller, 1986; Johnson et al., 2007; Weber et al., 2007). Across both studies, participants listed thoughts that did and did not align with their final answer, which suggests a degree of consideration and that the norm is not entirely pre-stored. The degree of thought clustering varied by norm—when SMRD was zero it indicated no thought clustering for randomness¹³. The topics that were most normative, according to prior literature (corporate fiduciary duty: McGowan, 2022; meal-choice: Wolfswinkel et al., 2024), exhibited the strongest clustering. In Study 1, the descriptive economic priorities norm and the injunctive environmental practices norm had the most clustered thoughts; and in Study 2, both norms in the meat frame were most clustered. This pattern of low-

¹³ SMRD was close to zero for three conditions—the injunctive economic priorities norm in Study 1 and both injunctive and descriptive vegetarian norms in Study 2.

clustering in counter-normative conditions is different between the individual choice and social cognition contexts. The classic clustering effect in query theory is demonstrated in the context of asymmetric discounting, where SMRD is positive for the accelerating choice frame and negative for the delaying choice frame (Weber et al., 2007). Another example is found in the context of the endowment effect, where SMRD is high and positive for the endowed condition and smaller yet still positive for the seller condition (Johnson et al., 2007). In neither case is SMRD very close to zero, as is observed in the counter-normative examples in the present chapter. This suggests that counterfactuals of norm conditions (i.e., counter-normative or uncommon practices) are not evaluated with the same attention as they are for individual choice. For descriptive norms, agreements between norm framing and reality appear to inform query order; and for injunctive norms, moral judgements or social desirability appear to inform query order.

While query features were associated with norm perceptions, the effect of frame and norm type conditions on perception were not consistently significant. In Study 1, query order predicted norm perceptions; and in Study 2, query order and proportion predicted perceptions of different norm topics. This cross-study difference is likely related to domain relevance and psychological distance, or the feeling of being removed from an object, person, or event by space, time or social boundaries. The corporate priorities and practices norm topics, which are abstract, are more distant than meal choice, which is a common and familiar type of experience for most people. Aligning with studies on psychological distance and conformity to the norm, we found that meal choice norms were less malleable than corporate environmental norms by the frame conditions. There is evidence that people conform more strongly to norms that are psychologically distant through the process of abstraction, which results in more reliance on group opinion (Ledgerwood & Callahan, 2012). In the meal choice scenario, participants were prompted to think very concretely which

affected the thought process but not, ultimately, norm perceptions. In the case where norms were more psychologically distant, norm perceptions were more sensitive to frame and norm type conditions. Greater psychological distance in Study 1 explains why norm perceptions were more malleable and, therefore, why query order was affected. Psychological distance may also affect the number of relevant memories that are cued, such that more distant norm topics have fewer personal memories and more proximal norm topics have more. This explanation is born out in the meal choice study: though the familiarity of the domain brings more experiences and hence more queries to draw on, this familiarity may also make norm perceptions less malleable. Thus, the frame can affect the thought process without changing the perceived norm. Taken together, these studies suggest construction of norm perceptions is more affected by contextual features when the domain is psychologically distant.

The exploratory findings from the memory coding analysis suggest that semantic memories dominate the evaluation of social norms in the thought listing paradigm¹⁴. The present findings on memory type are perhaps more aligned with the contemporary distinction between episodic and semantic memory, which suggests a more interconnected dynamic between the two (Renoult et al., 2019), where semantic memories often provide the structure or schema that supports episodic recall, especially in constructing future scenarios or recalling past events (Irish & Piguet, 2013). This memory framework suggests the construction of norm perceptions may cue the semantic schema without activating episodic recall.

Limitations and Future Directions

¹⁴ Although the prompt in Study 2 was designed to elicit more concrete memories by asking an estimation question in a familiar context, the majority of memories were semantic.

The main limitation of the present research is a question of directionality between the query metrics (order and content) and norm perception. Although the task is sequenced clearly to the participants, it is possible that participants answered the norm perception in their mind first and then justify their answers in the thought listing task. This possibility is consistent across all query listing conditions, so the comparison between conditions would not be affected. Additionally, the query metrics provide evidence that most participants consider both options, therefore the memory retrieval explanation of query theory is still relevant.

There is an underlying assumption in these studies that the thought-listing process during norm estimation is naturalistic and ecologically valid. I included a no-thought-listing control group in Study 1 to contextualize results from the agree and disagree frames. The norm perceptions from the control group had a similar pattern of norm perceptions—i.e., descriptive economic priorities and injunctive environmental practices were highest, followed by descriptive environmental practices and injunctive economic priorities was lowest. However, since the frames were different, this is not a true validation study. The thought-listing paradigm has been validated in the query theory literature in the context of individual choice (Weber et al., 2007). I include a full validation study in Chapter 4 of the present dissertation to evaluate the effect of the thought-listing paradigm on the relationship between social norm information and policy support. Taken together, there is evidence the thought-listing process does not affect the outcome measure but there is not a way to fully validate that thought content or order is not affected by the paradigm. However, there is no reason to think the effect of the paradigm would vary by condition, so between-condition comparisons are valid.

Future studies should utilize incentive-compatible designs. The present studies were not incentive compatible to reduce the interference with the naturalistic thought process of norm

estimation—in the real world, there is no incentive for accurate norm perceptions. However, including an incentive-compatible study to this line of work would offer insight into how motivating accuracy affects thought process and outcomes.

Finally, future work should investigate the role of affective amplification in norm construction (Kahneman & Miller, 1986). The literature suggests surprising exemplars receive more attention and might get more weight in construction; however, the present study was based on participant-generated memories and was not designed to measure this. Future investigation should measure the role of novelty and inconsistent evidence through this memory-retrieval lens.

Chapter 3: Social Norms in Media: How the Prominence of Social Norms in Cable Television Shapes Norm Perceptions and Attitudes about Democracy¹⁵

There is a general consensus that American democracy is under threat from political sectarianism (Finkel et al., 2020) and the weakening of democratic norms (Holliday et al., 2024). Empirical evidence finds that antidemocratic posturing is driven by political elites and has not, yet, spilled over to public attitudes (Holliday et al., 2024). However, high levels of affective polarization contribute to misperceptions about how much the public endorses antidemocratic norms (Iyengar et al., 2019; Lees & Cikara, 2020; Voelkl et al., 2024). While public attitudes toward antidemocratic norms are low, the public may hold misperceptions about other Americans and about the Democratic and Republican parties. This work investigates those social norm perceptions and the role of the information-environment in shaping misperceptions about antidemocratic actions.

The American public is immersed in an information-dense environment where signals about democratic norms come from elites and institutions, via traditional news media, and from known others and strangers, via direct observation and social media. This media is becoming increasingly polarized, particularly prime time television shows (Kim et al., 2022), which is linked to increased partisanship in viewers (Muisse et al., 2022) and increased engagement in public political discord (King et al., 2017). People, both knowingly and unknowingly, sort through the plethora of information to identify trends and patterns about the social world. Social norm perceptions shape the attitudes and behaviors that people think are permissible within a certain, e.g., democratic, system (Kalkstein et al., 2023).

¹⁵ This chapter was written with coauthors: Kris Nichols (Princeton University), and Elke U. Weber (Princeton University).

The ideology of different news sources has been shown to affect viewer's attitudes and increase partisan misperceptions of social norms (Sparkman et al., 2022; Pasket et al., 2022). While much of the scholarship has focused on consumer demand as the driver of political bias in media, recent evidence suggests that ideological slant also comes from the newsroom, the people creating the stories and framing (Hassell et al., 2022; Muise et al., 2022). This suggests that people are receiving partisan signals through the news even when they are not aware of it or seeking it out. I conceptualize news media as both a reflection of the public's opinions, priorities, and reactions and a source of elite and institutional signals. This dual function positions media as simultaneously mirroring societal views while also transmitting messages from institutions and those in power—these two are not necessarily aligned. The average news consumer is, therefore, exposed to true normative information about others alongside red herring information that exacerbates social norm misperceptions.

In this chapter, I leverage recent advances in large language models (LLMs) to investigate the prominence of social norm information in cable television news and, subsequently, how media diet affects attitudes and norm perceptions about democratic norm violation topics. First, I use a signal detection theory framework to evaluate how well different LLMs can identify norm information in a large-scale text dataset made up of transcripts from cable news (3,265 hours of television transcripts). Second, I present descriptive analyses of cable news coverage in the two months leading up to the 2024 Presidential Election. Third, I conduct a survey of registered voters (388 interviews) to understand how norm prominence within their media diet's coverage of democracy topics correlates with attitudes and norm perceptions about democracy.

Democratic norm violations

Recent history has illustrated the weakening of democratic norms in the United States. The contest over the outcome of the 2020 election and President Donald Trump’s “big lie” violated the normative peaceful transfer of powers following democratic elections, where the loser consents to the winner (Arceneaux & Truex, 2023; Clayton et al., 2021). Trump’s second term began with the most executive orders signed by a president on their first day in office (*Executive Order Data as of January 20, 2025*, 2025). These actions have been lauded by many Republican lawmakers (Broadwater, 2024; Kapur et al., 2025). The literature underscores the influence of such elite signals on the public (Clayton et al., 2021; Hall & Druckman, 2023; Constantino & Weber, 2021; Tankard & Paluck, 2017). This has led to many empirical investigations of public opinion, with some contradictory findings. A survey focused on the 2022 election found that the public had not adopted the same acceptance of norm violations as political elites, suggesting recent outbreaks of antidemocratic actions had not weakened public support of democracy (Holliday et al., 2024). A longer view of this dynamic, examining norm violations between 2006 and 2021, found that public endorsement of norm violations has been increasing amongst both parties (Littvay et al., 2024). This study also found that people were more supportive of democratic norm violations when their own party was in power. Beyond public attitudes, exposure to elite statements about antidemocratic actions weakens the public’s trust and confidence in those institutions (Clayton et al., 2021). This body of evidence tells us that the increase in elite signals that violate democratic norms are impacting public understanding of those norms, if not directly, by changing public trust.

I focus on five democratic norms from America’s Political Pulse, a longitudinal panel survey (Westwood & Lelkes, 2024). The five antidemocratic actions are: in-party politicians ignoring out-party judicial rulings, reducing polling stations in out-party-dominated areas, in-party use of executive orders to circumvent Congress, censoring media sources that spend more time

criticizing the in-party than the out-party, and loyalty to the in-party over the Constitution in the context of contested elections (Fasching et al., 2024; Holliday et al., 2024). Public endorsement is relatively low for each of the actions—the most-supported is the use of executive orders (supported by 41% of Democrats and 28% of Republicans), followed by media censorship (supported by 19% of Democrats and 24% of Republicans), party loyalty, ignoring court decisions, and reducing polling stations. Additionally, there is evidence of out-group misperceptions, where in-group members massively overestimate how much the out-group endorses the norm violations (Westwood & Lelkes, 2024). The high levels of out-group misperceptions could have significant long-term consequences for peoples’ trust in democracy and further increase affective polarization, which has been shown to erode socially beneficial coordination.

Social norm perceptions and elite cues

Much of the literature has focused on public opinion and behavior (King et al., 2017; Fasching et al., 2024; Holliday et al., 2024) and while this offers invaluable insight to understanding American democracy, I argue the role of social dynamics has been incompletely integrated. Empirical work has found that social norm messaging, through radio programs, can affect social norm perceptions without changing attitudes (Paluck, 2009). The present chapter integrates social norm perceptions about democratic norms with the existing opinion research because norm perceptions inform the long-term stability of social structures—thinking others support antidemocratic actions can weaken one’s personal support for democratic attitudes and behaviors. Self-categorization theory suggests that people conform to social norms to align with their in-group. Individuals who identify with an in-group adopt normative behaviors to minimize in-group heterogeneity and maximize intergroup differences (Terry & Hogg, 1996). This theory does not require the norm to be a true group preference, only that it is salient and characteristic of

the in-group (Pryor et al., 2019). This is particularly relevant in the media landscape—if news coverage presents antidemocratic actions as common or permissible by political elites in a party, viewers who identify with that political party may adopt those antidemocratic attitudes and behaviors.

This chapter focuses on individual psychology, in particular on how a person takes in and integrates signals about democracy (i.e., elite signals, peer opinions) into their mental representations of social norms. Elite cues, such as those communicated through media about political elites, may affect public opinion through two channels. First, political theory suggests that in highly competitive environments, such as American politics, elite endorsements have an outsized effect on public opinion and policy support (Druckman et al., 2013). Second, more recent work focused on the Supreme Court suggests that elite cues, when they run counter to known public opinion, have indirect and mixed effects on public opinion (Clark et al., 2023). In this case the role of the media is critical in shaping the public response to the conflicting elite cue and understanding about public opinion, this is to say, the media can moderate the amount and type of influence that elite cues have on public opinion. In the case of *Dobbs v. Jackson Women's Health Organization*, Clark and colleagues found that public opinion was not directly influenced, but social norm perceptions moved away from the Court ruling. The authors argue this was due to the media coverage following the ruling that highlighted public disapproval of the Court ruling and made salient true social norms and their divergence from the elite cue. In this chapter, I take on this theory of indirect effect in the context of democracy. Empirical evidence shows that political leaders are exhibiting more antidemocratic actions and posture, but the public has not followed, and the public does not endorse violations of democratic norms (Westwood & Lelke, 2024;

Fasching et al., 2024; Holliday et al., 2024). This suggests media coverage and reactions to antidemocratic actions are critical in shaping public attitudes and norm perceptions.

The effect of media diet on norm perceptions

Media and institutional influences play a powerful role in shaping social norm perceptions. Media outlets can facilitate both norm change and norm erosion through the content they disseminate (Gelfand et al., 2024). For instance, media portrayals such as radio soap operas have been found to influence societal norms by altering perceptions on issues like intergroup marriage (Paluck, 2009). In contrast, media exposure can also lead to norm erosion—witnessing norm violations or controversial behaviors in the media can reduce stigma and increase public willingness to express previously marginalized opinions, potentially eroding social norms over time. While digital devices are a prominent medium from which Americans get their news (58% of Americans report preferring to get their news from digital devices), the share of Americans who prefer getting their news from television is consistent over the past four years (33%), and a majority of Americans report getting their news from television at least sometimes (63%) (“News Platform Fact Sheet,” 2024). Moreover, many Americans turn to television for more than news; many turn to partisan networks with partisan pundits who unpack and inform on uncertain events. Opinion programs are typically organized around a prominent figure(s), whose political punditry is elevated, influencing their viewers’ behavior. For example, during the COVID-19 pandemic, areas where people watched more Hannity and Tucker Carlson Tonight had fewer pandemic precautions and high rates of infection and death (Bursztyrn et al., 2023). The prominence of partisan echo chambers in traditional media erodes norms differently for different partisan groups. Partisan channels tend to expose viewers primarily to ideologically congenial content, reinforcing pre-existing biases and polarizing opinions further. This effect is particularly pronounced in opinion

programs, where subjective commentary is often perceived as factual information. These findings reflect the still robust influence of television in shaping attitudes and behaviors, as well as the potential for television to facilitate elite cues.

Moreover, partisan media consumption is both highly selective and habitual. Viewers often gravitate toward channels that confirm their prior beliefs, resulting in limited exposure to cross-cutting perspectives. This self-selection, combined with the strong appeal of nonpolitical attributes such as engaging hosts and distinctive visual styles, further entrenches partisan echo chambers (Broockman & Kalla, 2024). Even weak partisans and independents may inadvertently become aligned with a channel's perspective, contributing to a concentrated audience that increasingly relies on partisan outlets for news. The motivations for turning to partisan television are multifaceted. While directional motives and a distrust of "straight news" drive some viewers, the entertainment value and accessibility of opinion programs also play crucial roles. These factors not only solidify habitual viewing patterns but also facilitate the framing of complex political issues in a manner that reinforces existing normative perceptions.

Media also influences social norm perceptions by sharing and amplifying institutional signals and elite cues. Institutional decisions—such as changes in textbooks or landmark judicial rulings—can change individuals' perceptions of social norms, even without shifting their personal beliefs. For example, Supreme Court rulings have been found to shift individuals' attitudes on the acceptability of gay marriage and abortion, though recent work has suggested that subsequent media coverage of these events play a pivotal role in how individuals unpack these events (Clark et al., 2023; Tankard & Paluck, 2017). For most Americans, one of the primary sources of information for unpacking these events is the media, and many turn to preferred partisan media sources for opinions they trust (Broockman & Kalla, 2024). This is particularly relevant to the

discussion of democracy, where there are many examples of political elites violating or endorsing the violation of democracy norms. In the past it has been difficult to estimate the effects of television due to a paucity of data and insufficient computational power. In the past 5 years, more research has emerged pointing out the unique effects of television consumption, in particular partisan television consumption, on partisan attitudes and perceptions (Kim et al., 2022). Given the advancement of these methods and the still sizable impact of traditional media in communicating social normative information, understanding the frequency with which norms are communicated in media and the effects of such communication remains an under-investigated area. In the present chapter, I investigate how one's media diet, specifically the way in which cable news channels cover violations of democratic norms, influences individual attitudes and norm perceptions on those same topics.

Leveraging LLMs to identify social norms

The rapid development and deployment of LLMs over the past decade have expanded both opportunities and uncertainties for large-scale text analysis in the social sciences (see Binz et al., 2025; Demszky et al., 2023 for critical perspectives). Traditional natural language processing (NLP) methods rely on dictionary-based search, identifying specific terms associated with a concept. Notably, words are evaluated without their surrounding context. A common example is emotion coding with Linguistic Inquiry and Word Count (LIWC), which has been widely used in and validated for psychological research (Tausczik & Pennebaker, 2010). However, dictionary-based methods are inappropriate and insufficient for tasks that require contextual understanding. Social norms are one such concept that require context to identify, a capacity that LLM classifiers provide. While manual annotation has been considered the gold standard for context-dependent classification, this approach often underestimates the risk of human coding errors.

LLMs offer a potentially scalable and accurate alternative for classification when properly validated and applied. Since no validated NLP method currently exists for identifying social norms in naturalistic text, this paper provides a thorough methodological evaluation. Accuracy represents the primary concern when using LLMs for large-scale coding tasks. To address this, I apply a signal detection theory framework and the relevant evaluation metrics to quantify this accuracy and compare the results between different LLM models (Lynn & Barrett, 2014; Demszky et al., 2023). Signal detection theory offers a more nuanced evaluation of accuracy than a single accuracy statistic, which is particularly important when comparing between multiple norm classifiers. Specifically, this framework deconstructs accuracy rates into hits, misses, false alarms, and correct rejections, which allows researchers to distinguish between type one and type two errors.

Chapter Overview

This chapter investigates how media coverage of democracy topics informs perceptions and attitudes about democracy. The chapter is divided into three sections. I first evaluate the appropriateness of using LLMs, specifically different versions of GPT-4, to identify norm information in naturalistic text. Next, I conduct a descriptive analysis of norm information in a large text dataset made up of cable news transcripts from September and October 2024. I evaluate the overall prominence of social norm information communicated through cable news and in particular the magnitude of norm information within democracy topic coverage. Sentiment analysis is integrated to investigate the overall valence (how positive or negative) of the coverage. Finally, I conduct a survey of frequent cable news viewers to probe the relationship between news coverage of democracy and viewers' attitudes and norm perceptions. I link participants' reported media diet from the survey (i.e., the shows and channels participants report watching most in the

prior month) to my text corpus to quantify their unique exposure to democracy coverage and the social norm prominence therein. I investigate whether norm prominence in one’s media diet is associated with attitudes and norm perceptions about democracy.

Methods

Television News Text Data

This chapter uses a text corpus of cable television news transcripts, scraped from FOX, MSNBC, and CNN in September and October 2024. The processed text is categorized across five key dimensions (Table 3.1): program metadata (channel, show, date, time), norm communication (presence of social norm information and type of norm information), democracy topic coverage (discussion of five democracy topics), sentiment (the degree of positive or negative language), and norm information about each democracy topic (the presence of norm information in coverage of the five democracy topics). Different classifiers are used for each dimension and are discussed in detail below.

Dimension group	Column	Values	Source
Program metadata	Channel	Fox, CNN, MSNBC	Stanford Cable TV News Analyzer
	Show	varies	
	Date	9/2024-10/2024	
	Time	00:00 – 24:00	
Communicates social norm info.	Contains norm	[0,1]	Fine-tuned GPT-4o
	Norm type	[1,2,3,4,5]	
Democracy topic is mentioned	Ignoring court decisions	[0,1]	Dictionary-based topic model (Table A3.6)
	Reducing polling stations	[0,1]	
	Use of executive orders	[0,1]	
	Media censorship	[0,1]	

	Party loyalty	[0,1]	
Sentiment	Sentiment	-1 – 1	Harvard-IV dictionary
Norm about democracy topic	Norm about ignoring court decisions topic	[0,1]	Contains norm x [Democracy topic]
	Norm about reducing polling stations topic	[0,1]	
	Norm about use of executive orders topic	[0,1]	
	Norm about media censorship topic	[0,1]	
	Norm about party loyalty topic	[0,1]	

Table 3.1. Summary table of the five dimensions that each text segments is coded for. Each text segment in the corpus is coded for each column from the noted source. The “communicates social norm information” dimension is coded using a range of sources for the method investigation and validation (including expert raters, novice rates, untuned GPT-4, untuned GPT-4o, ad fine-tuned GPT-4o) but the listed source here indicates the source used for the main analyses.

Text corpus. The corpus used for this chapter is made up of television news transcripts from the three largest news channels in the United States (MSNBC, FOX, and CNN). I extracted all transcripts from 9/1/2024 – 10/31/2024, across 24-hours of programming, from the Stanford Cable TV News Analyzer (Hong et al., 2021). The corpus includes a total of 3,265 hours of television transcripts from 126 unique television shows across the three channels. Due to data hacks during this timeframe some days have missing data (Bowman, 2024); all three channels are missing data from October 11, MSNBC is missing data from October 17 and October 25, and Fox News is missing data from October 12 through 14.

Text data processing. Transcripts were scraped by the Stanford Cable TV News Analyzer team and shared with us via email. The data is publicly available through a web app and smaller data

requests are managed through direct download. Commercials were not removed due to the technical limitations and resource constraints at the Internet Archive following a series of cyber-attacks in October 2024. Therefore, my data covers the entirety of what viewers were exposed to on cable news including programming and commercials. After processing the data, I randomly selected 1,000 segments for training and model evaluation, called the evaluation set.

Novice Rater Coding. 300 segments from the expert coded evaluation set were selected for an additional round of coding by novice raters. This random sample of segments were taken from September and October of 2024. The novice raters ($n = 266$) were recruited from Prolific. Upon entering the survey, participants were introduced to social norms, including definitions and examples. Social norms are introduced as “a behavior or attitude that is desirable or common among a group”. Examples are given of statements that do (e.g., “most people want a better future for their kids”) and do not (e.g., “today I went to the pool because it was very hot”). Raters are also told that, “Descriptive norms characterize what most people do (how things are), and injunctive norms characterize what most people approve or disapprove of (how things should be).” For each sentence, raters answer, “Does this sentence communicate a norm?” and answer with: 1 = no norm present, 2 = a descriptive norm is communicated, 3 = an injunctive norm is communicated, 4 = both a descriptive and injunctive norm are communicated, 5 = there is not enough information to determine.

Before qualifying for the full coding task, raters given a comprehension check made up of three text segments to code. Raters passed the check if they correctly answered 2 of the 3 coding questions. In the main survey, raters coded five randomly selected text segments from the sample of 300. For each segment, raters also reported their confidence in their rating as “low”, “medium”, or “high”.

Expert Coding. Each segment in the evaluation set ($n = 1,000$) was coded by two of the authors (JC and KN), viewed in a randomized order. Segments were coded for whether “the segment communicates a social norm”, with the same response options used by the novice raters. Segments were first coded individually by each author, then disagreements were reconciled through review and discussion.

Untuned GPT Procedure. The full evaluation set ($n = 1,000$) was coded by *untuned GPT-4* and *untuned GPT-4o* using the OpenAI API to query the gpt-4-0613 and gpt-4o-2024-08-06 models, respectively. The code for this query was written in R. I used a temperature of 0 to ensure the highest predictability, meaning if the model were run multiple times, variation in response would be minimal (Rathje et al., 2024). The classification tasks were run in January 2025.

Fine-tuned GPT-4o Procedure. I used the OpenAI online platform to fine-tune GPT-4o before running the model through the API classification task, described above. To fine-tune the model, I trained the GPT-4o base model on sets of 100 expert-coded segments. Training and testing data were formatted with a custom prompt, described below, and the corresponding expert code. Each training batch used 5 epochs and 5 batches to measure incremental change during tuning. Epochs determine how many times the model can pass through the training data and batches segment the training data into smaller sets. These parameters are set to the levels recommended by OpenAI. Tuning was evaluated using OpenAI’s built-in validation metrics (training loss and validation loss) and calculated using a separate testing dataset of 100 expert-coded segments. I ran two rounds of fine-tuning on two sets of training and testing data; the second round improved on the first model, as indicated by decreasing training and validation loss metrics. I achieved a final training loss of 0.29 and validation loss of 0.033, which suggests a good fit. The fine-tuned GPT-4o model was applied to the comparison set ($n = 300$) and the full September and October 2024 dataset ($n =$

158,179) by calling the OpenAI API through R, described above. An example of a segment that was classified as communicating social norm information comes from CNN on 9/14/2024: “...the vast majority of us, as Americans, have so much more in common than what separates us. Most Americans want a leader who brings us together as Americans and not someone who professes to be a leader who is trying to have us point our fingers at each other...”.

GPT-Prompts. To create comparable classifiers, all classifiers (i.e., GPT, expert coders, and novice coders) were given the same prompt. For each task, I used a tailored prompt that included the base question, an explanation of key concepts, and instructions for how to answer, followed by the text of the segment. The base prompt was: “Does the text below (directly or indirectly) communicate a norm? Descriptive norms characterize common behaviors or attitudes (how things are), and injunctive norms characterize what most people approve or disapprove of (how things should be). Answer only with a number: 1 if no norm is communicated, 2 if a descriptive norm is communicated, 3 if an injunctive norm is communicated, 4 if both a descriptive and injunctive norm are communicated, or 5 if there is not enough information to answer or the text is truncated. Here is the text:” The outcome of all coding was a number from 1-5. I also created a binary “is norm” code where 2-4 are coded as 1 for “is norm” and 1 and 5 are coded as 0 for “no norm”.

Performance evaluation metrics. The four classifier types (novice coders, *untuned GPT-4*, *untuned-GPT-4o*, and *fine-tuned GPT-4o*) are evaluated for accuracy compared to the expert-coded dataset. Two performance standards were evaluated: (a) whether the classifier accurately identified that a norm was communicated, using the binary “is norm” code; and (b) whether the classifier accurately identified the type of norm that was communicated, using the category code. In the language of signal detection theory, the signal was a norm segment, and noise was a no-norm segment; hits were defined as classifiers accurately identifying a signal, misses were defined as

failing to identify a signal, and false alarms were defined as inaccurately classifying a noise segment as a signal. For each classifier, hit rate and false alarm rate were calculated. Hit rate was defined as the number of hits divided by the number of signal segments. False alarm rate was defined as the number of false alarms divided by the number of noise segments. I computed nonparametric signal detection theory measures of sensitivity (A) and bias (b) to evaluate classifier accuracy. The nonparametric method is recommended for binary data, which I use in the present study (Zhang & Mueller, 2005). Sensitivity describes the classifier's ability to distinguish between signal and noise distributions; bias describes the classifier's liberal or conservative leaning (i.e., if the classifier is more likely to report signal or noise when there is uncertainty). Sensitivity and bias were calculated from hit and false alarm rates on the R-shiny Web App developed by Zhang and Mueller.

Sentiment Analysis. Sentiment analysis was conducted on the corpus using the Sentiment Analysis package in R. Through this package I calculated sentiment using the Harvard-IV dictionary (as used in the General Inquirer software), which matches words from each segment with a pre-validated list of positive and negative words. The positive score is then subtracted from the negative score to yield an overall sentiment per segment.

Democracy Topics. Given the richness, size, and topical heterogeneity of my corpus across channels and over time, these text data are well suited for a topic modeling investigation. In line with previous studies (e.g., Hunger & Paxton, 2022; Widmann & Wich, 2023; Eichstaedt et al., 2021), I employed a dictionary-based approach to assess the prevalence of texts related to five democratic norms. These norms were derived from five questions in the *America's Political Pulse* survey (Westwood & Lelkes, 2024). To identify these norms, I constructed a dictionary that treats each norm as a distinct topic and searched for corresponding keywords within each text segment

(full dictionaries in Table A3.6). A segment was coded as addressing a particular norm if it contained at least one keyword from the relevant topic dictionary. The dictionary was inspired by the Democratic Backsliding dictionary developed by Zhelyazkova and colleagues (2024) and was revised to suit the American context. The democracy topic classifier identifies news segments that are related to each of the five democracy topics, not just on the specific antidemocratic action. For example, the following segment from Fox News on 9/21/2024 discusses new polling station policies in Georgia and is flagged as broadly related to the polling station topic: "...from Georgia tonight, the state election board has just approved a new plan requiring hand counts for every ballot cast on election day. Three sworn officers at every polling station in the state will have to count the ballots by hand and make sure they match the machine count ...".

Norm Communication on Democracy Topics. To identify segments that communicate social norm information about one of the five democracy topics, the norm classifier and democracy topic classifier are combined. If both classifiers are positive (e.g., communicates a social norm and is about court rulings), then the segment is flagged. For example, this clips CNN communicates a growing public sentiment about the Supreme Court's overturning of *Roe v. Wade*: "... the issues this election cycle will impact women, like reproductive freedoms. And Harris emphasized that she believes more people support...abortion access since the Supreme Court overturned *Roe v. Wade*. I think that's also why in state after state, so many so-called red states and so-called blue states, when the issue has been on the ballot, the American people are voting for freedom...it's saying the government shouldn't be telling people what to do..."

Media Diet Study

The sample. To further investigate the effects of attributes of individuals' media diet on their attitudes and perceptions of norms, I recruited a sample of people who live in the United States

and regularly watch cable news. The target was to recruit 400 participants balanced on political party and sex (AsPredicted #197132) and the sample was collected on Prolific Academic in November 2024 prior to the 2024 Presidential Election. To qualify for the survey, participants had to report that they watched at least one cable news channel (Fox News, CNN, and MSNBC) at least once a week. After recruiting 725 participants into the survey, 335 were screened out for failing to meet the media diet requirement. As preregistered, participants who failed the attention check ($n = 1$) were excluded. After these exclusions, there were 388 US-based participants (46% female, average age = 41 years, 35% Democrat, 25% Independent, and 40% Republican).

Procedure. First, participants reported their media diet by answering how frequently they got news from various outlets in the last month, “never”, “every few weeks”, “once a week”, “every few days”, or “almost everyday or more” (Sparkman et al., 2022). Participants reported their media diet for the following news sources: Fox News, CNN, MSNBC, print media (New York Times, Huffington Post, or the Washington Post), national broadcast news (i.e., ABC, CBS, NBC, or similar), NPR (radio or online), conservative news outlets (i.e., Breitbart, Drudge Report, Newsmax, Rush Limbaugh, The Blaze, OAN, etc.), and liberal news outlets (i.e., Democracy Now, The Intercept, The Nation, Salon, Mother Jones, Common Dreams, etc.). Participants qualified for the study if they watch any of the three cable news channels (Fox News, CNN, and MSNBC) at least once a week.

Next, participants reported the three shows they watched most on cable news in the prior month. This was followed by the attention check, where one person was screened out. Then, participants were asked about their norm perceptions related to five antidemocratic topics—ignoring court decisions, reducing polling stations, the use of executive orders, media censorship, and party loyalty (Holliday et al., 2024). For each topic (in a randomized order), participants were

asked, “What percentage of Americans support the following statement:” and answered on a sliding scale from 0 to 100. The statement text is shown in Figure 3.4. Then, participants were asked for norm perceptions about Democrats and Republicans with, “What percentage of Republicans and Democrats support the following statement:” and answered on two sliding scales from 0 to 100, for each party (randomized order). Next, participants were asked to report their own attitudes about the five topics. Participants were asked “Do you agree or disagree with the following statement:” on a five-point Likert scale from “strongly disagree” to “strongly agree”. The norms and attitudes were derived from five democratic norm violation questions in the *America’s Political Pulse* survey (Fasching et al., 2024; Westwood & Lelkes, 2024). Finally, participants answered demographic questions about their age, gender, political party affiliation, political ideology, and the type of place they currently live (i.e., “big city”, “small city”, “suburban area”, “small town”, or “rural area”).

Measures. *Norm perceptions about antidemocratic actions* were calculated for each of three reference groups (Americans, Democrats, and Republicans). For each of the reference groups, participants were asked, “What percentage of [Americans/Democrats/Republicans] support the following statement:” on a scale from 0 to 100 for each of the five antidemocratic actions. For example, the statement for ignoring court decisions was, “If your party's president can't get cooperation from the other party's members of congress to pass new laws, your party's president should circumvent Congress and issue executive orders on their own to accomplish their priorities.” The full text for each statement is shown in the first column of Figure 3.4. The five-item battery was asked in reference to each group and covered support for ignoring court decisions, reducing polling stations, the use of executive orders, media censorship, and party loyalty (Cronbach’s α (Americans) = 0.79; Cronbach’s α (Democrats) = 0.87; Cronbach’s α (Republicans)

= 0.91). Following the methodology of Fasching and colleagues (2024), I created a composite measure for each reference group using a confirmatory factor analysis based on the unidimensional item response theory model.

Attitudes about antidemocratic actions were measured through a version of the same five antidemocratic items used in the norm perception questions. Following the same methodology, the five attitudes were combined (Cronbach's $\alpha = 0.80$) to create a composite measure.

Results

GPT as a tool to surface social norms from naturalistic text

Model comparisons. I evaluated the accuracy of three different GPT models as well as human raters in identifying social norm information from naturalistic text. The three models are GPT-4, GPT-4o, and fine-tuned GPT-4o. These models were selected because they cover the range of available and user-friendly models offered by OpenAI. Accuracy is defined as agreement with the expert classification. All evaluations are run on the same random sample of 300 transcript segments.

Using expert raters' ratings as ground truth, *human raters* were 64% accurate when identifying whether a segment contains norm information. Using a one-proportion z-test, I found human raters performed significantly better than chance ($X^2(1) = 24.0, p < .00001$). Human raters have a hit rate of 74% and a false alarm rate of 42%—indicating that human raters accurately identified 74% of the true norms and incorrectly identified 42% of the true non-norms as norms (Table 3.2). Next, I evaluated the three GPT models. The *GPT-4* model was 61% accurate overall, with a 76% hit rate and 48% false alarm rate. *GPT-4o*, the current flagship model, was 70%

accurate overall, with a 69% hit rate and 30% false alarm rate. Finally, the *fine-tuned GPT-4o* model was 73% accurate overall with a 43% hit rate and 8% false alarm rate.

To evaluate the sensitivity and bias of the different classifiers, I used Zhang & Mueller's (2005) non-parametric sensitivity (A), and bias is calculated as Zhang & Mueller's non-parametric bias (b). Sensitivity describes the standardized mean difference between the signal and noise distributions and ranges from 0.70 to 0.80 in this context (Table 3.2). Bias describes how conservative or liberal the classifier is with higher values indicating a more conservative criterion. The most liberal classifier was *GPT-4* ($b = 0.67$), reflecting its high false alarm rate. This model performed most similarly to *human raters*, who had a slightly more conservative bias ($b = 0.76$). *GPT-4o* was moderately conservative ($b = 1.02$) and *fine-tuned GPT-4o* was very conservative ($b = 2.21$).

The sensitivity and bias of these classifiers were used to determine the best model to surface social norm communication. Given the goal of this classifier is to identify norms from large text corpora that are prohibitively time consuming for manual coding, the best tool would be a conservative one to avoid inflating Type 1 errors. *Fine-tuned GPT-4o* was best fit because it had a high overall accuracy and hit rate, and a very low false alarm rate. This model is conservative which suggests high confidence that flagged norms are indeed norms. The tradeoff was the low hit rate, which suggests about half of the true norms would be missed. For the descriptive analysis, I run separate, parallel analyses with the *fine-tuned GPT-4o* and *untuned GPT-4o* models to balance the low hit rate and high overall accuracy of the fine-tuned model with the higher hit rate and lower overall accuracy of the untuned model. Compared to *human raters*, both *GPT-4o* models were more accurate overall and more conservative, meaning accuracy increased by hit rate going up and false alarm rate going down. The models were also significantly cheaper than *human raters*,

making these classifiers scalable, which was important for the scale of data in the present chapter. The two *GPT-4o* models outperformed *GPT-4* in overall accuracy and bias, due to *GPT-4*'s high false alarm rate.

Classifier	Overall Accuracy	Hit Rate	False Alarm Rate	Sensitivity (A)	Bias (beta) High = more conservative
Novice human raters (via Prolific)	64% ***	74%	42%	0.72	0.76
GPT-4 (untuned)	61% ***	76%	48%	0.70	0.67
GPT-4o (untuned)	70% ***	69%	30%	0.75	1.02
GPT-4o (fine-tuned)	73% ***	49%	9%	0.80	2.21
<i>Extended evaluation (n=1002)</i>					
GPT-4o	67% ***	67%	33%	0.73	1.00
GPT-4o (fine-tuned)	80% ***	56%	9%	0.83	2.03

Table 3.2. Summary of norm surfacing classifier accuracy. Four classifiers are evaluated for their accuracy compared to expert raters—novice human raters, untuned GPT-4, untuned GPT4o, and fine-tuned GPT-4o. Overall accuracy is the percentage of accurate classifications, regardless of type, and is compared to chance (50% accuracy) using a one proportion z-test, significance levels are shown as: *** $p < .0001$, ** $p < .001$, * $p < .01$, · $p < .05$. Sensitivity describes the standardized mean difference between the signal and noise distributions; bias describes how conservative or liberal the classifier is with higher values indicating a more conservative criterion.

Evaluating GPT-4o (untuned and fine-tuned) accuracy by norm types. After identifying that the *untuned and fine-tuned GPT4o* have the highest binary accuracy for identifying the

absence or presence of any norm, those two models were evaluated for their norm-type accuracy. Norm-type accuracy was evaluated as correctly categorizing each segment as “not communicating a norm”, “communicating a descriptive norm”, “communicating an injunctive norm”, or “communicating both a descriptive and injunctive norm”. Table A3.4 summarizes the model accuracy. The overall norm-type accuracy of *untuned GPT-4o* was 54%. *Untuned GPT-4o* had a higher hit rate for descriptive norms than for injunctive norms ($HR_{descriptive} = 38\%$, $HR_{injunctive} = 12\%$; $X^2(1) = 24.73$, $p < 0.0001$; Table A3.4). The model was less sensitive in norm-type accuracy evaluation than in the binary accuracy evaluation, suggesting the distribution between signal and noise distributions was smaller for the norm-type tasks. *Untuned GPT-4o* was also more conservative for the norm-type task than the binary task, as indicated by higher bias scores and lower hit and false alarm rates (Table A3.4). *Fine-tuned GPT-4o* was 74% accurate for the norm-type task. This model also had a higher hit rate for descriptive norms than for injunctive norms ($HR_{descriptive} = 51\%$, $HR_{injunctive} = 34\%$; $X^2(1) = 8.45$, $p = 0.0037$; Table A3.4). The *fine-tuned GPT-4o* model had similar sensitivity on the binary and norm-type tasks; like the *untuned* model, the *fine-tuned* mode had higher bias in the norm-type task. As was the case for the binary accuracy evaluation, *fine-tuned GPT-4o* was more accurate than *untuned GPT-4o* and is a better fit for the norm-type classification task.

Descriptive analyses of cable news data

Norm prominence and sentiment in cable news. Using the *fine-tuned GPT-4o* classifier, I find that nearly one quarter (24.5%) of text segments on cable news in September and October 2024 contained norm information ($SD = 0.43$). Norm prominence varied across channels (Table 3.3); MSNBC had the highest proportion of norm-segments (33.8%, $SD = 0.47$), followed by CNN (23.5%, $SD = 0.42$), and Fox News (17.1%, $SD = 0.38$). Using a three-sample test for equality of

proportions, I found these proportions differ significantly from each other ($X^2(2) = 3746.5$, $p < 0.001$). Sentiment was, on average, neutral—the corpus-level average was 0.058 (SD = 0.10), measured from -1 to 1 where negative numbers indicate negative, 0 indicates neutral, and positive numbers indicate positive.

Channel	Prop. Norms	Avg. Sentiment	Norm Avg. Sentiment	Not Norm Avg. Sentiment
MSNBC	33.8% (0.47)	0.055 (0.132)	0.067 (0.075)	0.048 (0.112)
CNN	23.5% (0.42)	0.062 (0.076)	0.067 (0.069)	0.060 (0.077)
FOX	17.1% (0.38)	0.057 (0.102)	0.061 (0.056)	0.056 (0.141)

Table 3.3. Summary of the proportion of norms and text sentiment by channel. For each channel, the proportion of text segments that communicate a norm, the average sentiment, and the average sentiment on norm and not norm segments are reported. Proportion is calculated as the number of text segments that communicate a norm divided by the number of segments. Sentiment scores are on a scale from -1 to 1 where negative numbers indicate negative sentiment, zero indicates neutral sentiment and positive numbers indicate positive sentiment.

Using a binomial logistic regression, I examined how norm prominence changed over time for each channel. Given the timeframe of this data, leading into the 2024 Presidential Election, it was important to understand whether base rates were changing. The norm presence (binary) classifier was regressed on channel, week, and sentiment with interactions between channel and week and channel and sentiment. The probability of text containing a norm increased over time on all channels (Figure 3.1a; Table A3.5). The rate of increase was slightly higher on CNN than on

MSNBC ($OR = 0.97, p < 0.001$) and Fox ($OR = 0.98, p = 0.011$). It is likely that increased election coverage was a driving factor.

Controlling for change over time and sentiment, I saw that the norm probability was not significantly different between CNN and Fox ($OR = 1.39, p = 0.21$) but MSNBC did have a significantly higher probability of containing a norm than CNN ($OR = 5.96, p < 0.001$). There was also a significant interaction between sentiment and channel, such that higher sentiment (more positive coverage) was associated with a higher probability of communicating a norm on MSNBC and CNN. The probability increased more per unit for MSNBC than for CNN ($OR = 2.92, p < 0.001$). For Fox, most positive sentiment was associated with less change in norm probability than for CNN ($OR = 0.44, p < 0.001$; Figure 3.1b).

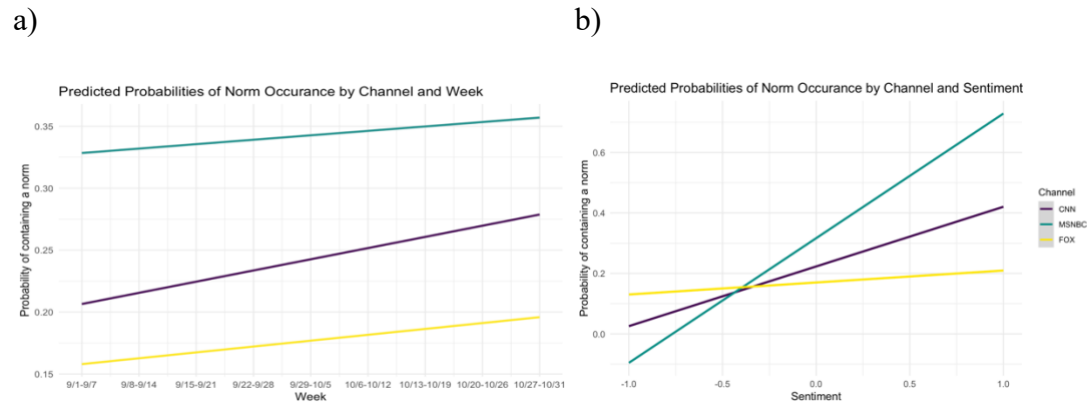


Figure 3.1. Interaction plots for a binomial logistic regression model of norm prominence in cable news transcripts from September and October 2024. Figure a shows the relationship between week (x-axis) and probability of text containing a norm (y-axis) varies by channel. Figure b shows the relationship between sentiment (x-axis) and probability of text containing a norm (y-axis) varies by channel. Color denotes channel– purple = CNN, teal = MSNBC, yellow = Fox. Full model results are shown in Table A3.5.

Descriptive norms were more common than injunctive norms in the cable news corpus; descriptive norms appeared in 14.2% of text segments and injunctive norms appeared in 10.2% ($X^2(1) = 1152.7, p < 0.001$). This pattern held for all three channels. On CNN 14.7% of segments contained descriptive norms and 8.8% contained injunctive norms ($X^2(1) = 1019.8, p < 0.001$). On Fox 10.4% of segments contained descriptive norms and 6.7% contained injunctive norms ($X^2(1) = 459.6, p < 0.001$). And on MSNBC 17.7% of segments contained descriptive norms and 16.0% contained injunctive norms ($X^2(1) = 45.43, p < 0.001$).

Coverage of democracy topics on cable news. To investigate the way democratic norm topics were covered on cable news, I applied a dictionary-based topic identifier to the corpus. A dictionary was constructed for each of the five democratic norm violations (use of executive orders, media censorship, party loyalty, ignoring court decisions, and reducing polling stations). Text segments were flagged as related to a democratic topic if they contained at least one relevant key-word for each topic. Note that my identification method flags text based on the broad topic, i.e., the use of executive orders, and not violations specifically. Overall, 23.8% (SD = 0.43) of text segments were about one of the democratic norm topics. This aligns with the literature that positions traditional media as a source of democratic and normative information to the public (i.e., King et al., 2017, Muise et al., 2022). The amount of coverage varied significantly across channels ($X^2(2) = 933, p < 0.001$); post-hoc pairwise comparisons with a Bonferroni adjustment confirmed all channels were significantly different from each other. MSNBC had the highest coverage of democratic norm topics (28.6%, SD = 0.45), followed by CNN (22.6%, SD = 0.42) and Fox News (20.7%, SD = 0.41). Using logistic mixed effects regression, I regressed topic coverage on channel and sentiment, interacted, with a random intercept for date. The full model results are shown in

Table A3.7. There was a significant interaction between sentiment and channel, such that coverage on CNN and Fox was more likely to be about democratic norms when there was negative sentiment. On MSNBC, text was more likely to be about democratic norms when sentiment was more positive. This suggests that Fox and CNN discussed negative aspects of democratic norms, such as examples of norm violations, while MSNBC was discussing positive examples of democratic norms. This suggests that media diet (i.e., the sources and quantity of news consumed) exposes people to different magnitudes of information about democratic norms. The channel-level differences in sentiment also suggest that channels cover democratic norms differently, which may also lead to partisan misperceptions.

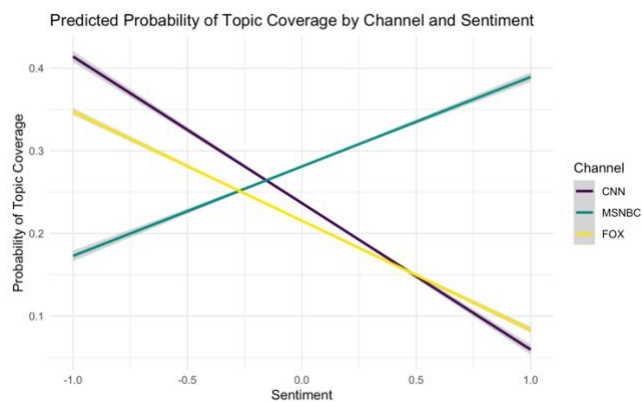


Figure 3.2. Interaction plot for the binomial logistic regression models predicting topic coverage in cable news transcripts. The plot shows the relationship between sentiment (x-axis) and probability of democratic norm topics being discussed (y-axis) varies by channel (color; purple = CNN, teal = MSNBC, yellow = Fox). The model includes a random effect for date. The full model results are shown in Table A3.7.

Within the coverage of democratic norm topics, 36.0% of text segments communicated a social norm ($SD = 0.48$), which is significantly higher than norm prominence across the full

transcript corpus ($X^2(1) = 2040, p < 0.001$). This suggests that the way democracy is covered on the news is characteristically different and contains a higher density of social norm information that does coverage on other topics. Norm prominence varied significantly across channels ($X^2(2) = 845, p < 0.001$); post-hoc pairwise comparisons with a Bonferroni adjustment confirmed all channels were significantly different from each other. The highest norm prominence was on MSNBC (44.7%, $SD = 0.50$), followed by CNN (34.5%, $SD = 0.48$), and Fox News (26.8%, $SD = 0.44$). To further examine how norm information was used when covering democratic norm topics, I used logistic mixed effects regressions. Norm presence was regressed on sentiment and channel, interacted, and a random intercept for date. Full regression results in Table A3.9. There was a significant interaction between channel and sentiment. On all three channels, more positive text segments were associated with higher norm prominence (Figure 3.3) and the association is weakest on Fox News.

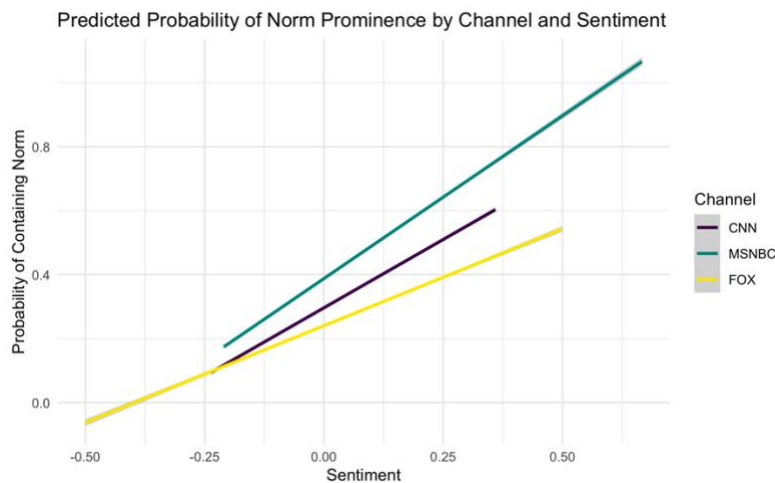


Figure 3.3. Interaction plot for the binomial logistic regression models predicting norm prominence in the coverage of democratic norms. The plot shows the relationship between sentiment (x-axis) and probability of containing a norm (y-axis) varies by channel (color; purple

= CNN, teal = MSNBC, yellow = Fox). The model includes a random effect for date. The full model results are shown in Table A3.9.

The relationship between media diet and attitudes and norm perceptions about democratic norm violations

Attitudes and norm perceptions about democratic norm violations. In a nationally representative sample by political party ($n = 388$), I measured attitudes and norm perceptions about the five democratic norm violations. Aligning with prior work, public endorsement of democratic norm violations was relatively low—only 9% of the sample endorsed the majority of democratic norm violations, i.e., endorsed violation of 3 or more of the 5 norms (Holliday et al., 2024). Endorsement of specific violations varied significantly by topic ($F(4,1389) = 61.53, p < 0.001$; generalized $\eta^2 = 0.065$), as indicated by a one-way repeated measures ANOVA test. Post-hoc paired analyses with a Bonferroni adjustment revealed that the use of executive orders was endorsed more than the other topics and reducing polling stations was opposed more than the other topics (p -values < 0.001). Norm perceptions were, notably, higher than personal endorsements; participants reported that, on average, 34% of Americans endorsed norm violations. This suggests a significant over estimation compared to true attitude levels. Norm perceptions also varied across domains ($F(4,1496) = 102.78, p < 0.001$; generalized $\eta^2 = 0.065$), indicated by a one-way repeated measures ANOVA . Post-hoc paired analyses with Bonferroni adjustments revealed the same pattern of results as the attitudes—that the use of executive orders is perceived as most-endorsed by the average American (41.8%, SE = 1.44) and reducing polling stations is perceived as least-endorsed by the average American (27.3%, SE = 1.45; p -values < 0.001).

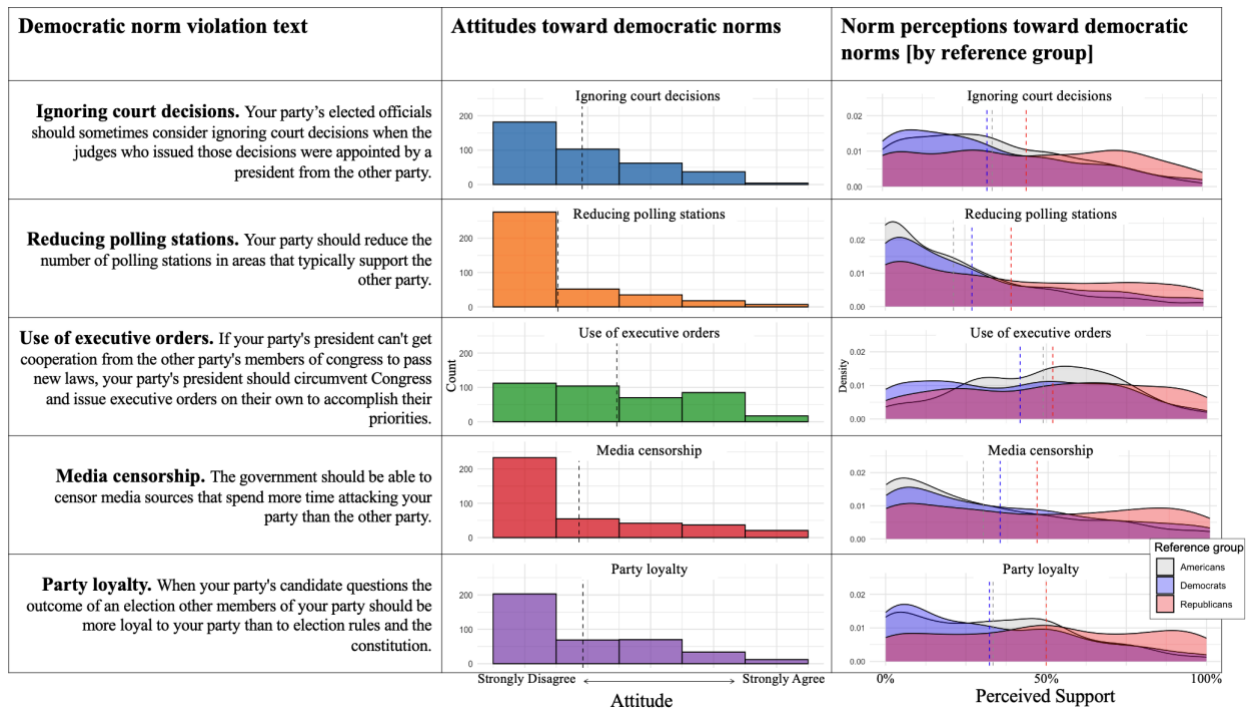


Figure 3.4. This figure shows the distribution of attitudes and norm perceptions about democratic norm violations from the media diet sample ($n = 388$). The first panel shows the text of each democratic norm violation. The second panel shows a histogram for each topic of participant attitudes endorsing the norm violation. Attitudes were measured on a 5-point Likert scale where lower numbers indicate opposition to antidemocratic actions and higher numbers indicate support for antidemocratic actions. The dashed line on each histogram shows the sample average. The third panel shows density plots of participants' norm perceptions about Americans (grey), Democrats (blue), and Republicans (red). The dashed lines show the sample means for each norm perception.

Show-level features of media diet. Next, I examined the effect of media diet on attitudes and norm perceptions. First, the amount of topic-relevant coverage, norm prominence in that coverage, and sentiment in that coverage was calculated based on the three shows participants watched the most over the prior month. For each participant, I average the topic coverage, norm prominence, and

sentiment for the three shows as an estimate of the characteristics of their media diet. Endorsement of democratic norm violations, perceived endorsement of democratic norm violations by Americans, perceived endorsement of democratic norm violations by Democrats, and perceived endorsement of democratic norm violations by Republicans were each regressed on topic coverage, norm prominence, and sentiment using linear regression analysis. Across the four models, the way norm violations were covered on cable news (measured as norm prominence and sentiment) had a more significant association with attitudes and norm perceptions than the amount of coverage (topic coverage).

As shown in Figure 3.5, topic coverage did not have a significant association in any of the models (all *p-values* > 0.10) while sentiment and norm prominence did. Attitudes about norm violations were significantly associated with norm prominence, such that more normative language in the coverage was associated with less endorsement of norm violations ($\beta_{Std.} = -0.24$, $t = -2.02$, $p = 0.044$). Interestingly, show-level media diet affected partisan norm perceptions but not the norm perceptions about Americans. Norm perceptions of Republicans (participants' perception of the percentage of Republicans who endorse norm violations) are significantly associated with the way norm violations were covered in cable news. Norm prominence is positively associated with perceived Republican norms, such that higher norm prominence in one's media diet was associated with a higher perceived Republican norm of endorsing democratic norm violations ($\beta_{Std.} = 0.38$, $t = 3.22$, $p = 0.0014$). Sentiment was negatively correlated with perceived Republican norms, meaning more negative coverage of democratic norm topics in one's media diet is associated with higher perceived Republican norms ($\beta_{Std.} = -0.27$, $t = -2.34$, $p = 0.020$). Norm prominence had a marginally significant association with perceived Democratic norms, such that lower norm

prominence was associated with higher levels of perceived Democrat norm endorsement ($\beta_{Std.} = -0.22, t = -1.88, p = 0.061$).

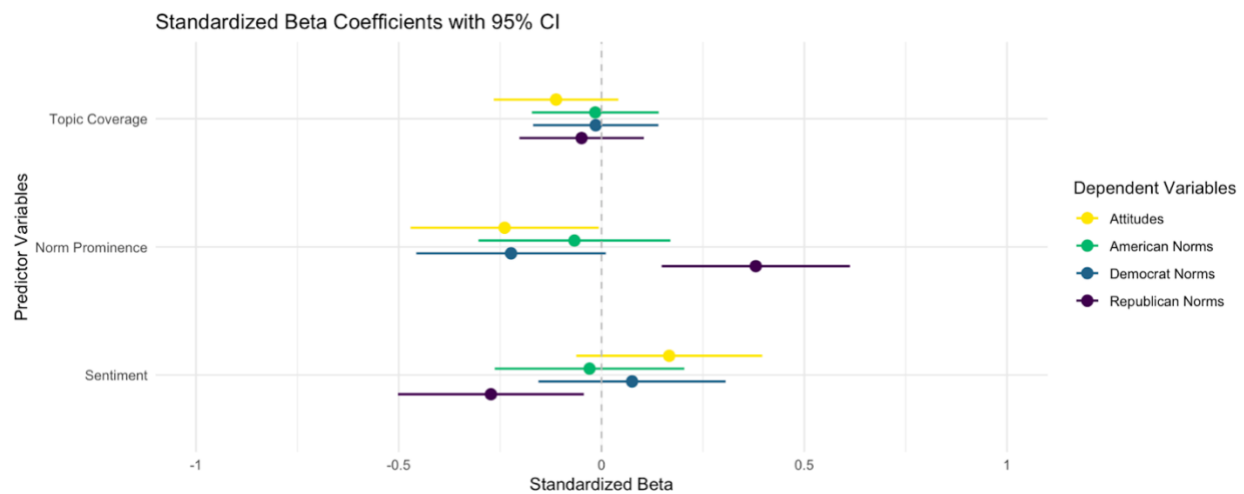


Figure 3.5. Standardized regression coefficients and 95% confidence intervals from four linear regression models predicting attitudes (yellow), norm perceptions of Americans (green), norm perceptions of Democrats (blue), and norm perceptions of Republicans (purple) toward democratic norm violations. Attitudes and norm perceptions were regressed separately on topic coverage, norm prominence, and sentiment in participants' most-watched cable news shows. The specific norm text is shown in Figure 3.4. Full model results are shown in Table A3.12.

Channel-level media diet and political polarization. To investigate the relationship between media diet and democratic norm violations beyond specific shows, I integrated channel-level news consumption and individual differences into the next analyses. In particular, I included the participants' political party and an interaction between political party and media diet in the regressions to understand partisan differences. I regressed attitudes and norm perceptions on channel-level news consumption and political party with interaction between media consumption and political party. Figure 3.6 shows the standardized regression coefficients; Table A3.13 shows

the full regression tables. For attitudes and norm perceptions of Democrats, I found a significant interaction between watching Fox News and being a Republican (Attitudes: $\beta_{Std.} = -0.51, t = -3.27, p = 0.001$; Democrat norms: $\beta_{Std.} = -0.33, t = -2.17, p = 0.031$). As shown in Figure 3.7, Democrats and Independents, but not Republicans, who watched more Fox News regularly reported higher personal endorsements of democratic norm violations and higher perceived norms of Democrats.

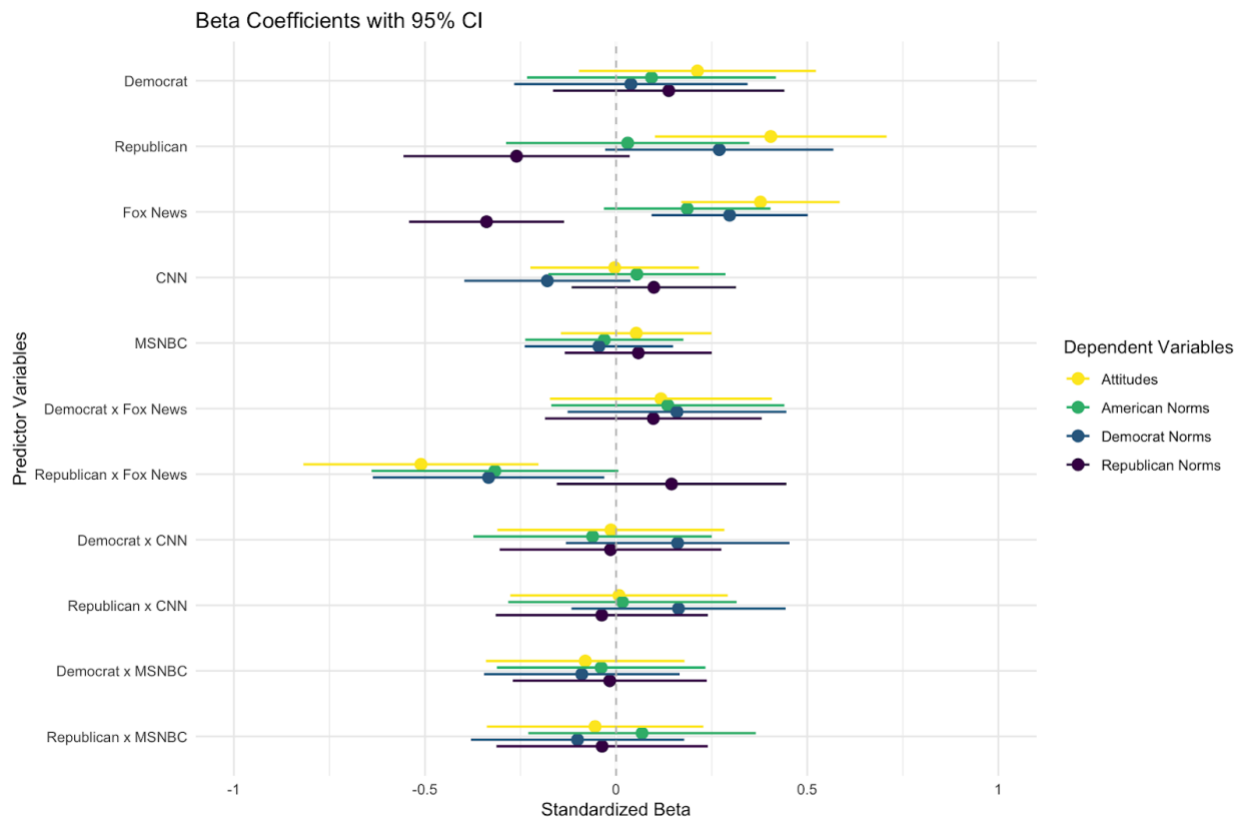


Figure 3.6. Standardized regression coefficients and 95% confidence intervals from four linear regression models predicting attitudes (yellow), norm perceptions of Americans (green), norm perceptions of Democrats (blue), and norm perceptions of Republicans (purple) toward democratic norm violations. Attitudes and norm perceptions were regressed separately on political party, amount of Fox News watched regularly, amount of CNN watched regularly, amount of MSNBC watched regularly, and the interaction between political party and each media diet metric. Political

parties are coded as binary variables. The specific norm text is shown in Figure 3.4. Full model results are shown in Table A3.13.

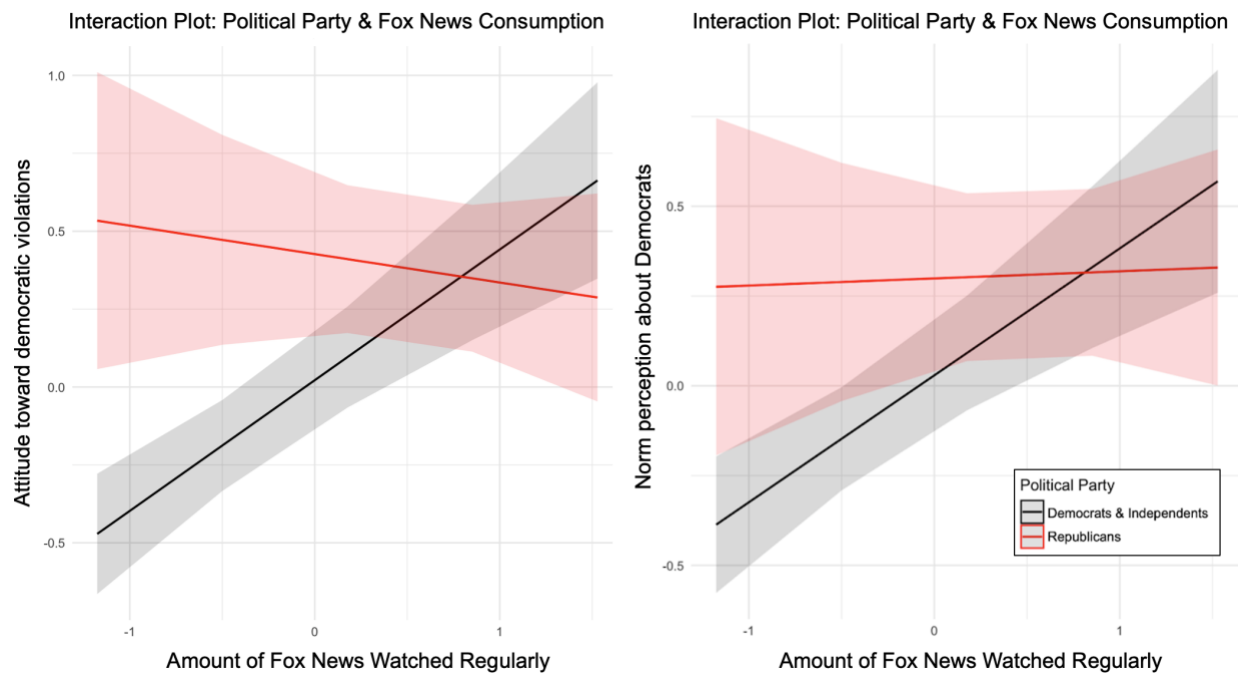


Figure 3.7. Interaction plots for the linear regression models predicting attitudes (left) and norm perceptions about Democrats (right), Figure 3.6. The plots show how the relationship between watching Fox News (x-axis) and attitudes (left, y-axis) or norm perception (right, y-axis) varies by political party (color; red = Republicans, black = Democrats & Independents).

Discussion

In this chapter, I leverage LLM tools to explore the prominence of norm information in cable news coverage on democracy topics and, subsequently, the effect on public perceptions. I find that a large portion of cable news coverage contains social norm information (36%) which is higher than the average normative-ness on the news. Coverage that contains social norm information is also more positive. The media diet study aligns with prior findings that while

individuals do not endorse democratic norm violations, they think others do (Holliday et al., 2024). The norm prominence within participants' media diets has asymmetric and partisan impacts on social norm perceptions.

Media is clearly a conduit of information that shapes partisan and, more broadly social, perceptions about democracy. This has been established across the social sciences. The findings extend this body of work by identifying and quantifying the social norm information within media coverage. Other researchers have investigated partisan bias (Muisse et al., 2022) and ideological skew (Hassell et al., 2022), but this is the first effort, to my knowledge, to specifically study prominence of social norms. I do not claim that social norm information is necessarily the driver of misperceptions, but I take up the argument of Clark and colleagues (2023) that media coverage of a controversial or antidemocratic action is consequential in shaping how people understand what is permissible in a democracy. And I offer a first estimation of and scalable method to identify social norm information within media coverage.

By pairing the survey data with the transcripts from participants' reported media diets, I can disentangle the effects of different features of media coverage on attitudes and norms. Social norm messaging, over and above pure information or news coverage, affects norm perceptions and attitudes. This underscores the importance of disentangling the type of social information people are exposed to. The results showed that exposure to social norm information about antidemocratic actions was associated with stronger opposition to norm violations but higher perceived Republican endorsement of those violations. While this partisan misperception finding is not new to the literature (Pasek et al., 2022), the association with quantified norm messaging is novel.

I leveraged channel-level analysis to investigate the sources of partisan asymmetry. Watching Fox News was more significant in shaping the attitudes and in-party perceptions of

Democrats and Independents. Drawing on learning theory, information on Fox News may be more impactful to Democrats and Independents because it is novel or “unconditioned” (Rescorla & Wagner, 1972). My findings suggest that the way Fox covers democracy is impactful to out-party members, perhaps because the information presented on the news interacts with what people experience in their personal lives. Further research is needed to disentangle the different effects that media and interpersonal signals have on partisan perceptions about democracy. The study offers preliminary evidence of partisan differences in the effect of social norm information on norms and attitudes about democracy.

Limitations and future directions

As with most archival analyses, there are data limitations that I must address. Although my data is extensive, the Internet Archive hack resulted in missing transcripts during the critical window of analysis. I addressed this by performing robustness analyses that exclude full days when a single channel is missing data. It is also important to note that I did not validate the accuracy of the transcribed television programs although the data has been used in peer reviewed papers (Kim et al., 2022; Hong et al., 2021). A limitation to the archival analysis is that the results from the dictionary-based method I use to identify “topic coverage” is prone to Type II errors. Although this is a standard method in the literature (e.g., Zhelyazkova et al., 2024), I cannot confirm that each mention of the terms is true coverage of democracy norms. The size of the corpus necessitates the use of this method. Finally, a limitation of the survey is that media exposure and norm perceptions are not synchronized. Survey participants report their prior month of media consumption on November 6 and the linked transcript data is from the month of October. Additionally, media diet in this survey was self-reported which is subject to error and limited to the three main cable news channels.

Future work should investigate the causal effect of social norm information in media coverage on attitudes and norm perceptions about democracy. The present study is correlational, relies on self-reported media diet, and does not use random assignment. While this study offers an important preliminary connection between the norm-coded television news corpus and norm perceptions, experimental work is needed to test the causality and direction of this relationship. Chapter 4 examines the effect of being exposed to social norm information on people's cognitive mechanism during policy evaluation.

Chapter 4: A process theory of social norm information interventions.¹⁶

People's behaviors, attitudes, and beliefs are shaped by the observed actions and perceived expectations of others. Psychologists have long acknowledged the effects of social norms (e.g., Sherif, 1937) and in the last half-century, social norms have been the subject of extensive research across the social sciences (Cialdini & Jacobson, 2021; Miller & Prentice, 2016). The predominant questions in the literature have focused on *why* social norms affect behaviors and attitudes, *which* norms (i.e., injunctive versus descriptive) are effective, and *when* they are impactful. The present chapter endeavors to explain *how* norms work through the lens of cognitive process tracing.

Effects of social norms

Social norms are those attitudes or behaviors that are desirable (injunctive norms) or common (descriptive norms) among a group or in a context (Cialdini & Goldstein, 2004; Legros & Cislighi, 2020). This definition of social norms is most aligned with the common definition in social psychology. Social norms are not individual preferences, attitudes, or habits; they are not instinctual or reactive behaviors; and they are not simply patterns in behavior (although some patterns are able to be attributed to social norms). Integrating the common definition in economics, social norms are those behaviors that are adopted because of a societal or communal expectation that this rule or practice should be followed (Bicchieri, 2006).

Social norms affect consequential individual behaviors and attitudes as well as collective phenomena. It is broadly accepted that social norms shape individual-level risky behaviors, such as smoking (Nyborg & Rege, 2003) and alcohol consumption (Burchell et al., 2013; Miller & Prentice, 2016). They also shape pro-social behaviors such as resource conservation (Abrahamse

¹⁶ This chapter was written with coauthors: Aya Salim (Princeton University), and Elke U. Weber (Princeton University).

& Steg, 2013), and how individuals interact with each other, for example bullying (Paluck et al., 2016) and beliefs about same-sex marriage (Tankard & Paluck, 2017). Social norms also inform what individuals expect from government such as the right to abortion (Clark et al., 2023), and civic engagement, like the motivation to vote (Gerber & Rogers, 2009).

Social norms about policy sit at the intersection of individual and collective solutions to societal challenges. Signals from the public shape governance and technological strategy in the energy transition (Otto et al., 2020). The public can communicate policy preference directly through public discourse, i.e., signing a petition or contacting a representative, or indirectly by voting for candidates aligned with their policy preferences. Policies also affect norm perceptions as institutional signals, through media discourse that adjust expectations of the beliefs, attitudes, or behaviors of others (Nyborg et al., 2016). The present research focuses on perceived norms that affect policy support, particularly related to the energy transition. New policies, specifically the Inflation Reduction Act (H.R. 5376, 2022) and the Infrastructure Investment and Jobs Act (H.R. 3684, 2021), have driven a rapid change in which new renewable energy technologies are economically feasible (Burgess et al., 2024), resulting in some uncertainty.

When public knowledge about new technologies is limited, the logic of plausible reasoning suggests that inference patterns become activated (Collins & Michalski, 1989). This indicates that individuals with less certain knowledge about emerging technologies rely more heavily on reasoning mechanisms—inductions, deductions, and analogies—that activate social norm cues from memory. These dynamics highlight the reciprocal relationship between policy initiatives and social norms by detailing how evolving public perceptions and meta-perceptions shape support for critical social and environmental policies. Through cognitive process tracing, this chapter

illuminates the mechanisms through which social norms operate to influence policy attitudes in contexts of technological and social change.

Why social norms affect behaviors and attitudes

Social norms affect behavior by shaping which behaviors and attitudes are considered acceptable or expected in a context. The focus theory of normative conduct (Cialdini et al., 1991) offers a meso-level explanation, arguing that norms guide behavior when they are made salient. From an evolutionary theory perspective, humans are born into cultures with complex systems of information that exceed individual capacity (Henrich, 2016). Humans have developed to learn social norms with a similar propensity as the learning of grammatical rules; such social learning has, therefore, emerged as an adaptive method to learn vast societal knowledge quickly and efficiently (Ostrom, 2000; Kendal et al., 2018).

The literature offers a few mechanism explanations of *why* social norms influence behavior—cognitive constraints, accuracy or efficiency, internalization, and conformity to avoid social sanction. The benefits of coordination and using norms as heuristics reinforces collective patterns of behavior. While direct experience is valuable, referencing others' experience or following their lead allows individuals to benefit from collective or dispersed knowledge. This is driven by the motivation to make “correct” or optimal decisions (Cialdini et al., 1990). Social norms can limit the range of behaviors that individuals consider, effectively excluding counter-normative actions from evaluation. This cognitive constraint means that people are less likely to consider behaviors that violate norms, which reduces the likelihood of such behaviors occurring (Kalkstein et al., 2023). Kalkstein and colleagues argue that a normative constraint is more effective at curtailing negative behaviors than personal motivation because norms reduce the

behavioral options that participants evaluate. However, norms can also entrench a harmful or anti-social habit (Andreoni et al., 2021), such as gendered discrimination (Becker, 2023).

Norms are internalized through processes like self-categorization, where individuals align their behavior with group norms to maintain a positive self-image and social identity (Pryor et al., 2019). This internalization makes norm-following part of one's self-concept, leading to adherence even in private settings where violations might not be detected (Gross & Vostroknutov, 2022). The focus theory of normative conduct proposes the last two mechanisms—social approval and accurate decision making (Cialdini et al., 1990). People follow norms to maintain a favorable social and self-image. The desire to be seen positively by others and oneself can drive adherence to norms, even when personal goals might conflict with them (Gross & Vostroknutov, 2022). The opposite is also true, namely people want to avoid being seen negatively and will adhere to a norm to avoid social sanction (Cialdini & Goldstein, 2004; Keizer & Schultz, 2012).

Which social norms affect behaviors and attitudes

Much of the social norms literature distinguishes between injunctive and descriptive norms, introduced by Cialdini and colleagues (1990). The common definition draws a clear distinction between the two—that injunctive norms describe approval or disapproval, and descriptive norms describe commonness or uncommonness. These two types of norms activate different motivational mechanisms to conform (Jacobson et al., 2011). Descriptive norms motivate through the desire for accuracy or to make the right decision; injunctive norms motivate through the desire for social approval and reciprocity. A meta-analytic review of the literature suggests that injunctive norm manipulations more strongly influence behavior, whereas descriptive norm manipulations have a greater impact on norm perceptions (Rhodes et al., 2020). However, the two are deeply interconnected—descriptive norms communicate injunctive information and vice versa

(Miller & Prentice, 2016; Blanton et al., 2008), a fact that may offer some explanation for why both types of norm interventions affect behavior (i.e., Bergquist et al., 2019).

Social norm information may also affect behaviors when a behavior is not currently a norm, through dynamic norm interventions (Sparkman & Walton 2017; Sparkman & Walton 2019; Mortensen et al., 2019). A dynamic norm describes the direction of change in others' behaviors or beliefs over time. These norms (e.g., people are taking more and more steps to reduce their carbon footprint) focus on the temporal dynamic and draw attention to the increase in behavioral and belief commonness rather than the base rate, or current level of commonness.

When social norms affect behaviors and attitudes

Across the behavioral sciences, researchers have run a large number of social norm interventions in just as many contexts. For example, in a meta-analysis of field experiments that use social norms to increase pro-environmental behavior the authors identified 74 studies with 91 independent field experiments and 227,730 total participants (Bergquist et al., 2019). Across this and other meta-analyses, researchers find high heterogeneity in the effectiveness of social norm interventions (e.g., Bergquist et al., 2019, Rhodes et al., 2020). Much of the heterogeneity in the literature is attributed to differences in operationalizing interventions and measurement. Moderation analysis identified that implicit norm interventions were more effective than explicit ones, as they led less psychological reactance and were more likely to be automatized (Bergquist et al., 2019).

The context of a norm intervention, i.e., the physical environment, the difficulty of changing behavior, and pre-intervention behavior, matter greatly in social norm influence. Environmental cues that provide physical evidence of the norm have been shown to increase intervention effectiveness (Cialdini et al., 1991). When a target behavior is challenging or incurs

discomfort, such as forgoing a plastic lid on one's coffee to-go cup, norm interventions tend to be less effective compared to behaviors with easy alternatives or modifications, like using only one paper towel when drying one's hands (Mundt et al., 2024). Furthermore, participants' pre-intervention behaviors are a significant determinant of intervention success, as evidenced by the OPOWER study, perhaps the most notable empirical example of social norm information affecting pro-environmental behavior (Allcott, 2011). In this field study, the average treatment effect of social norm information was near zero for the bottom two deciles of baseline household energy usage but exceeded 6% energy savings for those in the top decile.

Additional evidence suggests that observed information (i.e., memory content) influences the effectiveness of norm communication. When individuals hear that others support contributing to public goods (like community projects) but then see contradictory behavior (people not actually contributing), they can become resistant to contributing themselves (Miller & Prentice, 2016). However, there is an important exception: if someone was already inclined to contribute, then knowing their peers approve of such actions (injunctive norm support) can eliminate their fear of social disapproval. This effect occurs even when few people are actually contributing (descriptive norm). A person's starting point matters - peer approval can override concerns about going against the crowd, but primarily for those who were already predisposed toward the action in question (Smith & Louis, 2008; Smith et al., 2012).

In the context of pro-climate and environmental behaviors, it is particularly important to consider the political context of the intervention. There is empirical evidence of a partisan trust gap in climate science, such that conservatives have lower trust in science and scientists, with political consequences (Cologna & Siegrist, 2020; Azevedo & Jost, 2021). This has been tied to high levels of skepticism about the motivation of scientists and data accuracy. The moderating role

of skepticism on social norm interventions has not been empirically investigated, to my knowledge. The present chapter offers a first examination of whether and how social norm information activates climate science skepticism.

A query theory explanation of how social norms affect behavior

As discussed, the literature has offered broad evidence of when and why social norms affect behavior with less focus on the underlying cognitive mechanism, i.e., the question of *how* norms affect behavior. Query theory offers a process explanation of framing effects in individual decision making (Johnson et al., 2007; Weber et al., 2007; Composto et al. *under review*). Query theory says that, when faced with a judgment or decision, the order in which a person evaluates available choice options shifts the balance of support and, ultimately, helps to explain which option is chosen. One premise of query theory is that valuation questions are deconstructed into individual self-queries (Collins & Michalski, 1989). The way in which a question is framed shapes the decider's reasoning by influencing the order in which queries are posed and thus attended to (McKenzie & Nelson, 2003). Those queries activate memories that are related to the question topic and query (i.e., a query about reasons for choosing option A will activate memories about positive experiences with option A or negative experiences with other choice options). Finally, and importantly, the query that is posed first has a greater success in recall relevant memory content in the deliberation process, due to memory interference and inhibition where the attempt to recall one memory category temporarily suppresses access to other (competing) memory categories (Anderson et al., 1994).

The present chapter proposes a query theory explanation of social norm interventions. While norm interventions take many forms in the literature (Bergquist et al., 2019; Abrahamse & Steg, 2013; Miller & Prentice 2016; Rhodes et al., 2020), the present research focuses on norm

information interventions for their prevalence in the literature, similarity to framing effect interventions, and relevance to naturalistic information environments. The query theory explanation of social norm information interventions proposes that exposure to norm information affects the order and, thus, balance of participants' self-queries as they think through a decision. Norm information is assumed to affect query order by acting as a frame or prime: hearing that most people support a policy suggests that participants should begin by evaluating reasons to support the policy. As verified by dozens of query theory studies (Composto et al., *under review*), query order affects the balance of support, such that the option attended to first typically retrieves more supporting memories. Norm information might also affect query content by entering the thought process as a thought or piece of evidence in its own right. The thought-listing paradigm employed by query theory studies allows me to investigate these two alternative mechanisms, which are not mutually exclusive.

The present chapter explores how providing social norm information affects policy support by shaping the automatic thought process through which people evaluate evidence that supports or opposes that policy. Specifically, when norm information is presented, it draws attention to the normative option first (whether that be supporting or opposing the policy) by affecting the order and content of self-queries, which ultimately affects policy support. I present evidence that when social norm information supports (versus opposes) a policy, it increases policy support by making thoughts aligned with the norm information more accessible and prominent in people's minds.

Chapter Overview

I leverage the naturalistic type-aloud procedure (Johnson et al., 2007; Weber et al., 2007), to identify which cognitive mechanism(s) explain the effect of norm information interventions in

five studies. Understanding the mechanism(s) that underpin these interventions is crucial to designing impactful and scalable interventions. In Study 1, I tested the query theory explanation of social norm interventions in a hypothetical policy context with both a positive and negative norm provision treatment group. In Studies 2a-c, I tested the hypotheses in the context of real existing policies and norm information. Each of these studies employed a different policy and the norm information provided accurate information; therefore, there was only a single norm intervention (describing the actual support) for each policy. Study 2a focused on an energy efficiency policy with a positive norm treatment, Study 2b focused on a labor organizing policy with a positive norm treatment, and Study 2c focused on an energy technology policy with a negative norm treatment. Using the same policy and norm context as Study 2c, Study 3 manipulated query order to test the causality of the query process theory explanation. Next, Study 4 extended the findings into a more naturalistic information environment by providing policy context and norm information through real television news clips. The policy topic for Study 4 was support for large-scale solar farms and there were three conditions: a support norm treatment, an oppose norm treatment, and a no-norm control. Finally, I conduct a validation study to determine whether the type-aloud procedure accurately captures the naturalistic thought processes people experience when deciding whether to support specific policies.

I hypothesize that norm interventions will first direct attention to the status quo (as communicated through the norm), influencing the query order and content, and leading to policy support that is more aligned with the norm information (compared to a non-norm control condition). The following hypotheses are tested:

- *Hypothesis 1*: Social norm information affects policy support—positive norm increases policy support and negative norm decreases policy support. (Studies 1, 2a-c, and 4).

- *Hypothesis 2*: Social norm information influences the query order (H2a) and query content (H2b) of participants' thought process—positive norms lead to early support thought clustering and a greater proportion of support thoughts; negative norms lead to earlier oppose thought clustering and a greater proportion of oppose thoughts. (Studies 1, 2a-c, and 4).
- *Hypothesis 3*: Query order (H3a) and query content (H3b) mediate the effect of social norm information provision on policy support (Studies 1, 2a-c, and 4).
- *Hypothesis 4*: When query order is explicitly manipulated (soliciting it in the “natural” order, which was hypothesized and identified in Studies 2a-c, versus the opposite or “unnatural” order), the effect of norm information on policy support (H4a) and query content (H4b) is attenuated (Study 3).

Study 1: Effect of social norm information interventions on hypothetical policy decisions

Method

In this online study, I provided a first test of the cognitive mechanisms of social norm interventions on policy decisions in a hypothetical context. Participants were shown a hypothetical policy decision scenario and then were randomly assigned to a positive social norm treatment, a negative social norm treatment, or a no norm control condition. Participants were asked to type out the thoughts that went through their head as they evaluated the policy and then reported their support of the policy. Finally, participants were shown each thought they had listed previously and asked to categorize the thought as an argument that either increased or decreased policy support.

Participants

To detect a small-to-medium effect size ($f = 0.15$) at 80% power as preregistered (AsPredicted #149978), our target was to recruit 440 participants. I collected a sample on Prolific Academic in November 2023 by advertising a survey experiment to respondents located in the United States. I stopped data collection once 486 responses had been completed. As preregistered, I excluded participants who failed one or more of the four attention checks ($n = 30$) and participants who spent more than 60 minutes on the survey ($n = 1$). Of the 486 responses that were collected, I excluded 31 and Prolific excluded 14 for missing IDs or demographic information. Exclusions did not differ by condition (control versus positive norm versus negative norm), $\chi^2(2) = 0.0012$, $p = 0.99$. After these exclusions, there were 441 US-based participants (40 % female, 74 % white, 57 % four-year college degree or more; median age group = 45-55 years, median income group = \$50,000-74,999, 47% Democrat, 27% Independent, and 22% Republican).

Procedure

After giving informed consent, participants underwent two training tasks to learn how to use the online platform and paradigm, followed by an attention check. All participants then read a realistic, hypothetical scenario about a decision being made by the local municipal planning committee in their community about whether to add a new crosswalk to a highly-trafficked road (Table A4.1). Participants in the positive norm condition then saw an infographic saying “80% of people in your neighborhood support adding a crosswalk” ($n = 147$), participants in the negative norm condition saw an infographic saying “80% of people in your neighborhood oppose adding a crosswalk” ($n = 148$), and control participants did not receive any information about the neighborhood’s opinion ($n = 146$; Table A4.1). The norm treatments were framed as survey findings reported by a local newspaper and were accompanied by a pie chart graphic to communicate the proportion of support versus opposition. Next, all participants were asked to “list

the thoughts that go through your mind as you evaluate if you support or oppose putting in a new crosswalk” (Table A4.1). Participants typed one thought at a time into a free-text entry box and could enter as many thoughts as they wanted; a new box would appear after hitting return. After they indicated that they were done with their thought listing, participants were asked if they supported or opposed building a new crosswalk (*binary decision*) and how much they supported or opposed building a new crosswalk on a 5-point Likert scale (*strongly support, somewhat support, neither support nor oppose, somewhat oppose, strongly oppose*). Participants answered a willingness to pay question on the same 5-point Likert scale: “The installation would cost you approximately [\$50/\$100/\$130] and the maintenance would require [\$25/\$50/\$65] annually. Would you approve of such a policy knowing what it would cost you personally?” The indicated pairs of cost values were randomly assigned.

After answering the dependent variable questions, participants were shown each thought they had listed in the order in which they had entered them. For each thought, participants answered, “Was this thought in favor of increasing or decreasing support for the policy?” with options “it was in favor of increasing support”, “it was in favor of decreasing support”, “it was in favor of both increasing and decreasing support” or “it was in favor of neither”. For each thought, participants also answered, “Was this thought a reason to pay more or less for the policy?”, with options “it was a reason to pay more”, “it was a reason to pay less”, “it was both a reason to pay more and to pay less” or “it was a reason for neither”. After all thoughts were categorized, participants answered a second pair of attention checks and demographic questions. Participants reported their gender, age, race, education, income, and political party affiliation (see Table A4.2 for detailed demographic questions).

At the end of the survey, participants in the positive or negative norm treatment group were asked about their skepticism toward the norm information. Participants were re-shown the norm infographic and answered, “To what extent were you skeptical about the accuracy of this information?” with options, “not at all skeptical”, “slightly skeptical”, “somewhat skeptical”, and “very skeptical”.

Results

I replicated prior findings that social norm information affects policy support, in particular that positive norm information increases, and negative norm information decreases support for a policy (de Groot & Schuitema, 2012; Bolsen et al., 2014). As shown in the top panel of Figure 4.1, the supportive social norm information increased policy support compared to control ($M_{\text{support norm}} = 3.23$, $SD_{\text{support norm}} = 1.16$; $M_{\text{control}} = 2.90$, $SD_{\text{control}} = 1.35$; $t(284) = 2.23$, $p = 0.027$) and the opposing norm information decreased policy support ($M_{\text{oppose norm}} = 2.45$, $SD_{\text{oppose norm}} = 1.47$; $t(290) = 2.74$, $p = 0.0064$). Policy support was significantly higher in the support norm condition than in the oppose norm condition ($t(279) = 5.05$, $p < 0.0001$)¹⁷. This supports Hypothesis 1.

On average, participants listed 4.00 thoughts ($SD = 1.75$). Participants in the control condition listed more thoughts than those in the two treatment conditions ($M_{\text{control}} = 4.29$, $SD_{\text{control}} = 1.73$; $M_{\text{support norm}} = 3.86$, $SD_{\text{support norm}} = 1.61$; $M_{\text{oppose norm}} = 3.86$, $SD_{\text{oppose norm}} = 1.86$). Using two-sample t-tests I compared the number of thoughts listed in the control condition to the support norm condition ($t(289) = 2.17$, $p = 0.031$) and the oppose norm condition ($t(291) = 2.02$, $p = 0.045$) and found both had significantly fewer thoughts, although the absolute difference was small. There

¹⁷ Willingness to pay (WTP) also varied significantly across conditions, as indicated by a two-way ANOVA controlling for amount ($F(2, 432) = 4.43$, $p = 0.013$). WTP was significantly higher in the support norm condition than in the oppose norm condition ($t(293) = 2.98$, $p = 0.003$).

was no significant difference in the number of aspects listed between the two norm conditions ($p=0.996$).

Next, I examined the query metrics, order and content. Query order, measured as standardized median rank difference (SMRD)¹⁸, was positive in all three conditions which indicates that reasons to support the policy were generated first. Positive support thoughts clustered earlier in the support norm condition than in the control condition; positive support thought also clustered earlier in the support norm and control conditions than the oppose norm condition ($F(2,438)=5.03$, $p=0.0070$). The SMRD difference between the support and oppose norm conditions was significant ($t(283) = 3.68$, $p = 0.0017$), shown in the middle panel of Figure 4.1. The negative norm weakened the initial focus on reasons to support the policy, supporting Hypothesis 2a. Evaluating the balance of support thoughts in the aspect list, the pattern was similar ($F(2,438)=6.73$, $p=0.0013$), shown in the bottom panel of Figure 4.1. The proportion of support thoughts was greater in the support norm condition than in the oppose norm condition ($t(286) = 3.67$, $p = 0.00030$). This supports Hypothesis 2b. These results support the crucial QT prediction that changes in query order (SMRD) will result in a different frequency of support versus oppose thoughts and thence the balance of policy support.

¹⁸ Order is measured as the standardized median rank differences (SMRD) between arguments to “support the policy” and arguments to “oppose the policy”. SMRD is defined as $2(MR_i - MR_p)/n$, where MR_i is the median rank of arguments to “support” in a respondent’s sequence of thought listing preceding the choice, MR_p is the median rank of arguments to “oppose”, and n is the total number of thoughts listed (Johnson et al., 2007). SMRD takes the value of +1 if all arguments to “support” are recalled first, and -1 if all arguments to “oppose” are recalled first, and 0 if the two types of arguments do not cluster but are randomly interspersed.

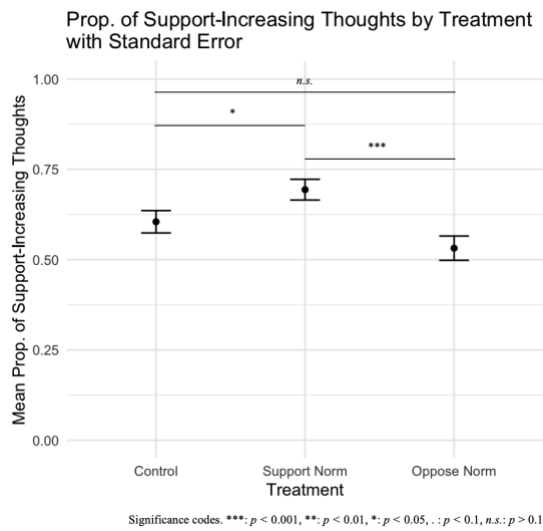
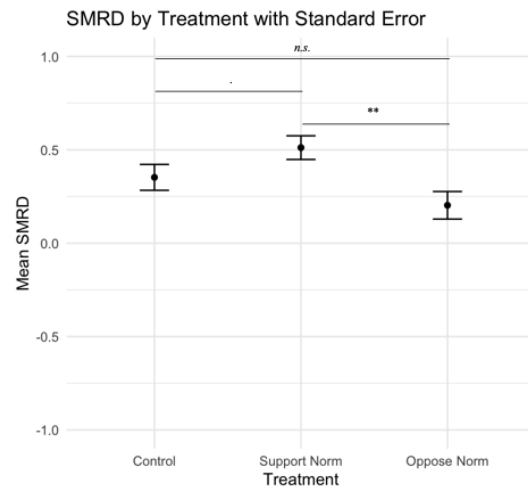
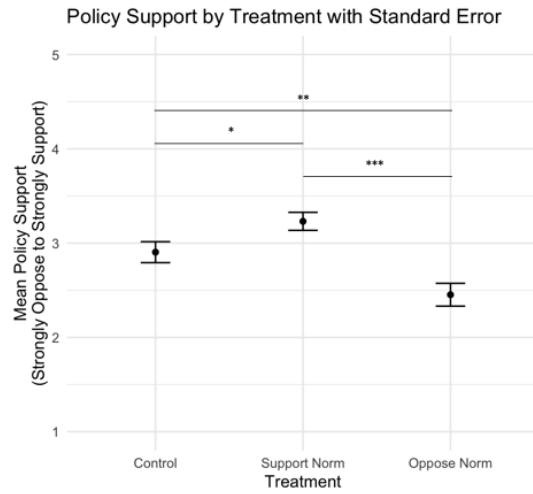


Figure 4.1. Results of Study 1: policy support (top panel), standardized median rank difference (SMRD; middle panel), and proportion of support thoughts (bottom panel) are reported by condition. Policy support is measured on a 5-point Likert scale where 5= strongly support, 3 = neither support nor oppose, and 1= strongly oppose. SMRD is measured on a scale from -1 to 1 where positive values indicate support-thoughts clustered earlier than oppose-thoughts, negative values indicate the opposite, and a value of zero indicates randomly dispersed thoughts. Proportion of support thoughts is measured as the proportion of thoughts listed that are in support of the policy, out of the total number of thoughts.

To examine the query theory mechanisms, I ran two mediation analyses. The analyses were conducted using the “mediation” package in R, with 10,000 bootstrap resamples to compute confidence intervals for the indirect effect (Tingley et al., 2014). First, I examined whether query order mediated the effect of norm treatment on policy support. The results suggested that SMRD was a partial mediator, as indicated by the significant indirect effect ($\beta = 0.29$, 95%CI [0.11,0.48], $p = 0.0012$). The direct effect remained significant after accounting for the mediator ($\beta = 0.49$, 95%CI [0.23,0.74], $p = 0.0002$), suggesting partial mediation. SMRD mediated 37% of the effect of norm information on policy support (95%CI [0.17,0.61]; Table A4.3). Second, I examined whether query content mediated the effect of norm treatment (support versus oppose norm) on policy support. The results suggested that the effect of norm information on policy support was partially mediated by the proportion of support thoughts, as indicated by the significant indirect effect ($\beta = 0.42$, 95%CI [0.19,0.65], $p = 0.0002$). The direct effect remained significant after accounting for the mediator ($\beta = 0.36$, 95%CI [0.14,0.58], $p = 0.0012$), suggesting partial mediation. Proportion of support thoughts mediated 54% of the effect of norm information on policy support

(95%CI [0.32,0.77]; Table A4.4). As predicted by query theory, the two query mechanisms were correlated but not redundant ($r = .79$), so I ran a third mediation analysis on a combined query mechanism factor and found the combined query mechanism partially mediated the effect of norm information on policy support¹⁹. The query metrics also mediated the effect of social norm information on willingness to pay²⁰. Taken together, the mediation analyses support Hypothesis 3.

Discussion

Study 1 found preliminary support for the query theory process explanation of social norm interventions. In the hypothetical policy scenario, the effect of social norm information on policy support was mediated by both query order and content. Norm information in support of the policy led to earlier clustering of supporting thoughts, compared to norm information in opposition to the policy. The earlier clustering of support thoughts was associated with a greater proportion of policy supporting thoughts, which offers the first evidence of a query theory mechanism in the context of social cognition.

Study 2 a-c: Effect of social norm information interventions on real policy decisions

Method

¹⁹ Query content (proportion of support thoughts) and order (SMRD) are combined into a single prominence-of-support-thoughts factor using a principal components factor analysis (Weber et al., 2007). Mediation analysis found that prominence-of-support-thoughts partially mediates the effect of norm information on policy support, as indicated by the significant indirect effect ($\beta = 0.39$, 95%CI [0.18,0.62], $p < 0.0001$). The direct effect remained significant after accounting for the mediator ($\beta = 0.39$, 95%CI [0.16,0.62], $p = 0.0012$), suggesting partial mediation. Prominence-of-support-thoughts mediated 50% of the effect of norm information on policy support (95%CI [0.28,0.74]; Table A4.5).

²⁰ SMRD fully mediated the relationship between social norm information and WTP, as indicated by the significant indirect effect ($\beta = 0.25$, 95%CI [0.09,0.43], $p = 0.002$) and the non-significant direct effect after accounting for the mediator ($\beta = 0.27$, 95%CI [-0.05,0.58], $p = 0.10$). Proportion of support thoughts also fully mediated the relationship between social norm information and WTP, as indicated by the significant indirect effect ($\beta = 0.37$, 95%CI [0.17,0.59], $p = 0.0002$) and the non-significant direct effect after accounting for the mediator ($\beta = 0.15$, 95%CI [-0.14,0.44], $p = 0.33$).

To examine the cognitive mechanism of social norm interventions for non-hypothetical policy decisions, U.S.-based participants were recruited to deliberate on three existing policies. Building on the design of Study 1, I designed three policy stimuli based on policies in place or proposed in the U.S. and corresponding social norm information from nationally representative polls. As in Study 1, participants first read information about a policy topic before being shown either current social norm information (treatment) or neutral information (control). They then listed their thoughts, reported policy support, and categorized their thoughts. The three policy topics are: national energy efficiency standards on major appliances and consumer products (Study 2a), national protections for employees to engage in collective bargaining and wage negotiations (Study 2b), and national investment in carbon capture and storage (CCS) programs (Study 2c). The corresponding social norm treatments in Study 2a-b (i.e., existing levels of support for each policy) were positive norms and the treatment in Study 2c was a negative norm. The specific policy context and associated treatment stimuli are shown in Table A4.6.

Participants

The samples for Study 2a and 2b were collected together in October 2023 via Prolific Academic. To detect a small effect size of norm information on policy support ($d = 0.25$) at 80% power as preregistered, the target sample size was 450 participants, solicited by advertising a survey experiment to respondents located in the United States (AsPredicted #146450). Participants were excluded if they failed one or more of the four attention checks or if their completion time was greater or less than 3 SD above/below the sample mean completion time. Of the 468 responses that were collected, 31 were excluded. Exclusions did not differ by condition (treatment versus control x policy topic, $\chi^2(3) = 0.029$, $p = 0.99$). After these preregistered exclusions, I achieved 437 US-based participants (45% female, 55% four-year college degree or more; mean age group

was 35-44 years). The distribution of political affiliation was 48% Democrat, 32% Independent, 13% Republican, and 7% other.

The Study 2c sample was collected in July 2024 via Prolific Academic. To detect a small-to-medium effect size of norm information on policy support ($d = 0.30$) at 80% power as preregistered, the target sample size was 350 participants, solicited by advertising a survey experiment to respondents located in the United States (AsPredicted #183536). Participants were excluded if they failed one or more of the four attention checks or their completion time exceeded 60 minutes. Of the 366 responses that were collected, 11 were excluded. Exclusions did not differ by condition (treatment versus control, $\chi^2(1) = 0.0028$, $p = 0.96$). After these preregistered exclusions, I achieved 355 US-based participants (50% female, 54% four-year college degree or more; mean age group was 45-54 years). The distribution of political affiliation was 34% Democrat, 37% Independent, 28% Republican, and 1% other.

Procedure

Studies 2a-c followed the same procedure as Study 1. Participants read a policy scenario, then treatment participants (but not controls) were shown social norm information, then all participants listed their thoughts, answered the policy support and willingness-to-pay dependent variable questions, categorized their thoughts, and answered demographic and skepticism questions. The norm information for participants in Study 2a was, “A strong majority of registered voters in the United States think that global warming should be a high priority for the president and Congress. ... A strong majority of Americans support setting stronger energy efficiency standards for new buildings, appliances, and cars/trucks/SUVs.” The norm information for participants in Study 2b was, “A strong majority of Americans support labor unions and most support workers going on strike to negotiate for better wages, benefits, and working conditions. A

strong majority of Americans think it is bad for the United States that labor unions have gotten smaller.” The text norm interventions in Studies 2a and 2b were accompanied by a graphic that summarized real survey findings about the norm described. Study 2c used a short animation to illustrate the norm information. The norm information text in Study 2c was, “... you will view a short animation that visualizes distrust in the local governments that are responsible for regulating and permitting new CCS projects. 10,000 people answered: do you trust or distrust the local governments that are responsible for regulating and permitting new CCS projects? Each box represents the opinion of 100 individuals that were surveyed— green means trust and red means distrust. 61% of Americans distrust the local governments that are responsible for regulating and permitting new CCS projects.”

The only difference from the Study 1 procedure was the use of new stimuli (shown in Table A4.6) and different cost amounts in the willingness-to-pay question. In the CCS study (Study 2c), a new policy skepticism question was also added. The willingness-to-pay question (if “passing this policy would cost you approximately [\$50/\$10/\$1] per month (equivalent to a [\$600/\$120/\$12] increase in what you owe on your annual income taxes, would you approve of such a policy knowing what it would cost you personally?”) was answered on the same 5-point Likert scale. The pairs of values were randomly assigned; this question was asked in Studies 2a-b and not in Study 2c. In addition to the norm skepticism question, I asked participants in the CCS policy condition, “to what extent are you skeptical about the accuracy of [the policy context] information?” with options, “not at all skeptical”, “slightly skeptical”, “somewhat skeptical”, and “very skeptical”.

Results

I replicated the typical effect of norm information on policy support for all participants in the CCS policy scenario (Study 2c) and for those participants who were not skeptical of the

provided norm information in the energy efficiency and labor protections policy scenarios (Studies 2a-b). To test for the effect of norm information on policy support, one-way ANOVAs were run separately for each study. There was not a significant effect of positive norm information on policy support in Study 2a ($F(1,216) = 0.27, p = 0.60$) nor in Study 2b ($F(1,217) = 0.33, p = 0.57$). When skeptics, i.e., participants who were ‘skeptical’ or ‘very skeptical’ of the norm information provided in the treatment conditions, were removed, however, there was a significant effect on climate policy support, Study 2a, ($F(1,172) = 4.20, p = 0.042$; Figure 4.2a) but not on union policy support, Study 2b ($F(1,199) = 1.15, p = 0.29$; Figure 4.2b). There were 44 skeptics in the climate policy scenario and 18 in the union policy scenario; a test of the equality of two proportions indicated that the proportion of skeptics was significantly higher in the climate conditions than in the union condition ($\chi^2(1) = 13.41, p = 0.0025$). A three-proportion test of equality indicated that the number of skeptics in the climate policy condition varied by political party, with independents being the least skeptical ($\chi^2(2) = 7.65, p = 0.022$). The proportion of skeptics did not vary by political party in the union policy condition ($\chi^2(2) = 1.75, p = 0.42$). The thoughts listed by skeptical participants suggested participants came in with strong negative prior opinions about the policy and immediately rejected the norm information and disputed the accuracy of the stimuli. For non-skeptics, positive norm information increased climate policy compared to control ($t(159) = 2.14, p = 0.034$, Figure 4.2a).

In Study 2c, negative norm information reduced CCS policy support compared to the no information control ($t(348) = -3.13, p = 0.0019$, Figure 4.2c). This result was robust to skepticism, subgroup analyses of skeptics and non-skeptics both showed significant treatment effects on policy support. Taken together, there is asymmetric evidence to support Hypothesis 1, such that the effect

of negative norm information about CCS was larger and more homogeneous than the positive norm intervention on energy-efficiency standards.

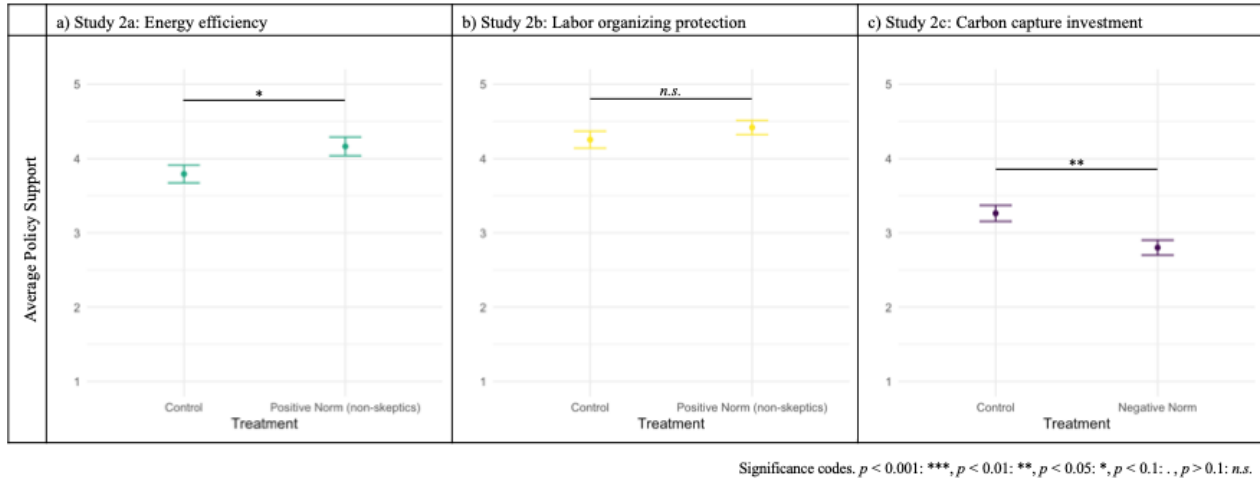


Figure 4.2a-c. Results of Study 2a-2c: policy support is reported by condition for each study. The treatment condition for Studies 2a and 2b was a policy-supporting norm information and the treatment condition for Study 2c was a policy-opposing norm information. For Studies 2a and 2b, only non-skeptics are plotted for the treatment condition. Policy support is measured on a 5-point Likert scale where 5= strongly support, 3 = neither support nor oppose, and 1= strongly oppose.

Participants listed an average of 4.20 thoughts ($SD = 1.78$) during the policy deliberation phase. The number of thoughts listed does not vary by treatment in Study 2a ($F(1, 216) = 1.30, p = 0.26$), Study 2b ($F(1, 217) = 0.12, p = 0.73$), nor Study 2c ($F(1, 353) = 1.75, p = 0.19$). To examine variation in number of thoughts listed by policy across Studies 2a-c, a two-way ANOVA was run on the number of thoughts with policy and treatment interacted ($F(2, 786) = 3.83, p = 0.022$). Participants listed significantly fewer thoughts in the energy-efficiency scenario ($M = 3.93, SD = 1.63$) than in the unions ($M = 4.24, SD = 1.90$) or CCS ($M = 4.34, SD = 1.79$) scenarios ($b = -0.42, SE = 0.15, t = -2.74, p = 0.0063$).

Studies 2a-2b: Query theory metrics

Examining the query theory mechanisms, thoughts listed during policy consideration were affected by social norm information, as predicted in Hypothesis 2, but the specific mechanism differed between the positive and negative norm treatments. For non-skeptics in the positive norm conditions (Studies 2a-b) SMRD was significantly more positive for the treatment than control condition, as indicated by a two-way ANOVA controlling for policy ($F(1,371) = 4.20, p = 0.041$). SMRD was significantly higher for the labor union policy than for the energy efficiency policy ($F(1,371) = 7.50, p = 0.0065$) and there was no significant interaction between treatment and policy²¹. However, SMRD was positive for both conditions across both policies, suggesting that searching for reasons to support the policy was the default predisposition even in the control group. Next, I examined the effect of positive norm information on balance of support using a two-way ANOVA with policy topic and treatment. The balance of support thoughts listed by non-skeptics only differed marginally, though in the predicted direction ($F(1,371) = 2.84, p = 0.093$; $EMM_{\text{support norm}} = 0.73, SE_{\text{support norm}} = 0.033$; $EMM_{\text{control}} = 0.66, SE_{\text{control}} = 0.028$; $t(349) = 1.51, p = 0.13$).

As in Study 1, a mediation analysis was run to investigate the effect of the query theory mechanisms. For non-skeptics, I examined whether query order mediated the effect of positive norm information (versus control) on policy support in the energy efficiency and labor protections scenarios. The analysis controlled for political party, policy topic, and the interaction between treatment and policy topic. The effect of positive norm information was partially mediated by SMRD. There was a significant indirect effect of treatment on SMRD ($\beta = 0.13, 95\%CI [0.0045, 0.26], p = 0.041$). The direct effect remained marginally significant after accounting for the

²¹ Post-hoc comparison of the estimated marginal means (EMMs) revealed that the EMM for the positive norm treatment group was marginally significantly higher than the EMM for the control group ($t(371) = 1.86, p = 0.064$). This indicated that support thoughts were listed earlier in the positive norm condition than in control.

mediator and the bootstrapped 95% confidence interval included 0, suggesting partial mediation ($\beta = 0.18$, 95%CI [-0.00064, 0.37], $p = 0.051$): SMRD mediated 41% of the positive norm information intervention on policy support (Table A4.8).

Study 2c: Query theory metrics

In the CCS policy scenario, the negative norm information had only a marginal effect on query order (Figure 4.3a). Using a one-way ANOVA, I found a marginally significant effect of condition on SMRD ($F(1,353) = 2.77$, $p = 0.097$; $EMM_{\text{oppose norm}} = -0.15$, $SE_{\text{oppose norm}} = 0.068$; $EMM_{\text{control}} = 0.016$, $SE_{\text{control}} = 0.071$; $t(342) = 1.66$, $p = 0.097$). Notably, SMRD was near zero in both conditions, indicating thoughts were randomly interspersed, likely due to participants' unfamiliarity with the policy topic²². The negative norm condition did have a significant effect on the balance of support thoughts compared to control ($F(1,353) = 6.85$, $p = 0.0092$). Post-hoc comparison of the estimated marginal means (EMMs) revealed a large decrease in balance of support thoughts in the oppose norm condition compared to the no information control ($EMM_{\text{oppose norm}} = 0.39$, $SE_{\text{oppose norm}} = 0.031$; $EMM_{\text{control}} = 0.032$; $t(353) = -2.62$, $p = 0.0092$; Figure 4.3b).

A mediation analysis was conducted to examine whether query content mediates the effect of negative norm information (versus control) on policy support in the CCS context. The results suggested that the effect of negative norm information on CCS policy support was mediated by balance of support thoughts, as indicated by the significant indirect effect ($\beta = -0.31$, 95%CI [-0.53,-0.08], $p = 0.0086$). The direct effect was not significant after accounting for the mediator ($\beta = -0.16$, 95%CI [-0.34,0.02], $p = 0.087$), suggesting full mediation. Proportion of support thoughts

²² Chapter 2, Study 1 of the present dissertation surveyed a nationally representative sample ($n = 1,935$) about carbon capture and storage technologies, finding only 30% of respondents had heard of these technologies.

mediated 66% of the effect of negative norm information on policy support (95%CI [0.29,1.08]; Table A4.7).

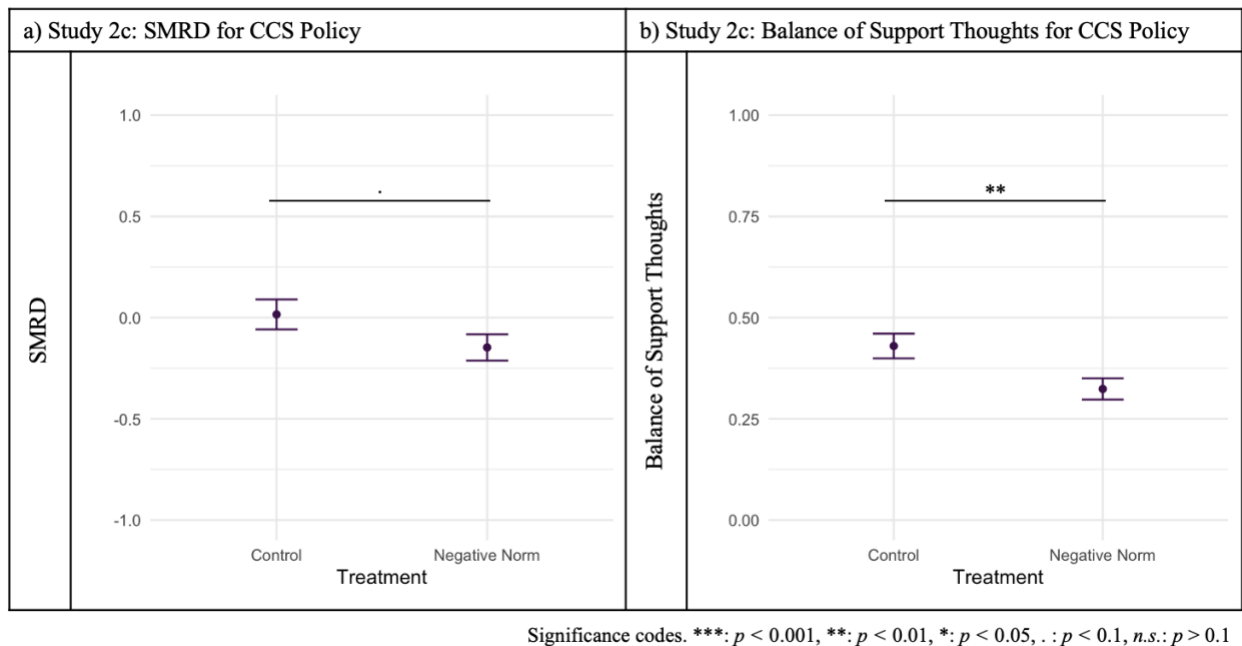


Figure 4.3. Results of Study 2c: standardized median rank difference (SMRD; left panel), and balance of support thoughts (right panel) are reported by condition. SMRD is measured on a scale from -1 to 1 where positive values indicate support-thoughts clustered earlier than oppose-thoughts, negative values indicate the opposite, and zero indicates randomly dispersed thoughts. Proportion of support thoughts is measured as the proportion of thoughts listed that are in support of the policy, out of the total number of thoughts.

Discussion

Taken together, Studies 2a-2c support the broad hypothesis that the queries automatically issued to evaluate policy support are affected by the provision of norm information. The asymmetric effect of positive vs. negative social norm information on query theory metrics and

subsequent mediated policy support aligns with prior literature that finds negative norm information to be more impactful than positive norm information (Bolsen et al., 2014). In Studies 2a-b, I found that query order partially mediated the effect of positive norm information on policy support for participants who were not skeptical of the norm information. In Study 2c, I found that query content fully mediated the effect of negative norm information on policy support in the full sample. One notable difference between the policies is that the policy topics in Studies 2a-b were more familiar to participants than the Study 2c policy topic. CCS is a largely unknown energy technology (Pianta et al., 2021), which may explain why the query order was close to zero, suggesting randomly interspersed thoughts. Query order for the control conditions of Study 2a and 2b were significantly greater than zero, which suggests that the natural query order for considering these policies even without any norm information is to first consider arguments for supporting them. Since these studies used actual policies and actual norm information, I cannot disentangle whether it is the norm direction (positive versus negative norm information) that affects the relative strength of differences in query order versus content and/or whether it is the content of the policies, perhaps their familiarity, that is the driver. I address this in Study 4 where I designed positive and negative norm treatments for the same actual policy context.

Study 3: Causal investigation of query order mechanism

Method

The first two sets of studies provide evidence for a query theory explanation for the effect of social norms on policy decisions through cognitive process tracing, with correlational evidence utilized in the mediation analyses described for these studies. In Study 3 I tested the causal nature of query order in the mediation of the effect of social norm interventions on policy decisions. Using

the CCS policy stimuli introduced in Study 2c (Table A4.6), I manipulated query order during the thought listing phase. In Study 2c, I found that the natural query order for the CCS policy is mildly positive in the control condition (support thoughts come first in the thought list) and negative in the negative social norm treatment condition (oppose thoughts come first in the thought list). Study 3 employs a 2x2 between-subjects design, where participants are randomly assigned to either the negative norm treatment or the no-norm control condition (the same stimuli as used in Study 2c), crossed with a different prompted and explicit order in which support and oppose thoughts are elicited, either in the “natural” order (as observed in Study 2c for the norm treatment or control condition) or in the opposite “unnatural” order.

Participants

To detect a medium-to-small effect size ($d = 0.30$) of norm condition on policy support within each order condition, at 80% power as preregistered (AsPredicted #183536), our target was to recruit 700 participants. We collected our sample on Prolific Academic in July 2024 by advertising a survey experiment to respondents located in the United States. I excluded participants who failed one or more of the attention checks ($n = 91$) and participants who spent more than 60 minutes on the survey ($n = 2$). Participants were also excluded for failing to follow order instructions ($n = 79$), i.e., if participants in the support first condition listed all oppose thoughts before listing any support thoughts or vice versa. Exclusions did not differ by condition (treatment versus control, $\chi^2(1) = 0.0029$, $p = 0.86$; by thought listing condition, $\chi^2(1) = 1.98$, $p = 0.16$). After these exclusions, I achieved 522 US-based participants (53% female, 80% white, 58% four-year college degree or more; mean age group = 45-55 years, median income group = \$50,000-74,999, 30% Democrats, 37% Independents, 29% Republicans). Given the high number of

exclusions, I ran a sensitivity analysis and found there was 80% power to detect a medium-to-small effect size of a two-sided t-test in both order conditions ($d_{\text{unnatural}} = 0.36$, $d_{\text{natural}} = 0.34$).

Procedure

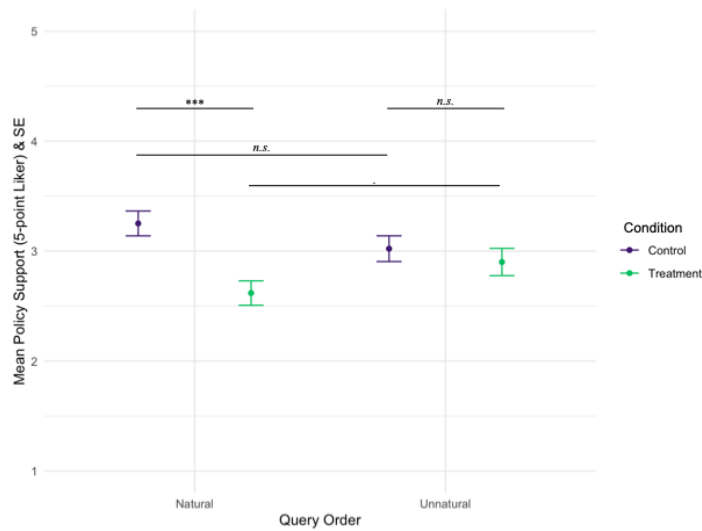
This study follows the procedure used for Study 2c, the CCS policy scenario, except for the thought-listing task. In Study 3, participants were prompted separately to list (a) reasons to support the policy and (b) reasons to oppose the policy. Half of the participants received the two queries in the order hypothesized to be natural for their condition as indicated by the results of Study 2c. The natural order in the negative norm condition was to list reasons to oppose the policy before listing reasons to support the policy; the natural order in the no-norm control condition was the reverse. The other half of participants received the two query prompts in the opposite, unnatural order.

Results

Figure 4.4a shows that when support and oppose thoughts were explicitly prompted in their natural order, the effect of norm information on policy support in Study 2c was replicated. Policy support is lower in the negative norm information condition than in the control condition ($t(272) = -3.99$, $p < 0.0001$)²³. When the two types of thoughts were prompted in the opposite, unnatural order, there was no significant difference in policy support between the treatment and control conditions ($t(239) = -0.71$, $p = 0.48$).

²³ As predicted by query theory, the effect of the negative norm information was larger in Study 3, where order was explicitly manipulated, than in Study 2c, where it was not ($t(348) = 3.13$, $p = 0.0019$). The same was true of the effect of condition on proportion of support thoughts.

Figure 4.4b shows that the query-order manipulation also affected the balance of support thoughts participants listed²⁴. Consistent with the observed policy support and as predicted in Hypothesis 4, the proportion of support thoughts differed significantly by norm information condition when queries were in the natural order, ($t(272) = -4.74, p < 0.0001$), but not when they were in the unnatural order, ($t(244) = -0.89, p = 0.38$).



²⁴ The number of thoughts listed did not vary by treatment, query order, or the interaction of the two ($F(1,517) = 1.30, p = 0.26$; $F(1,517) = 0.035, p = 0.85$; $F(1,517) = 0.20, p = 0.66$, respectively).

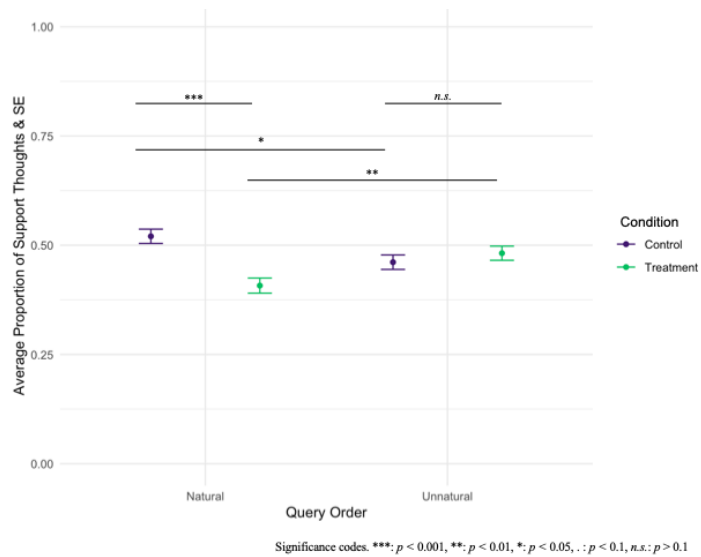


Figure 4.4. Results of Study 3: policy support (top panel) and proportion of support thoughts (bottom panel) are reported by condition and query order. Condition is noted by color: purple indicates control and green indicates treatment. Policy support is measured on a 5-point Likert scale where 5= strongly support, 3 = neither support nor oppose, and 1= strongly oppose. Proportion of support thoughts is measured as the proportion of thoughts listed that are in support of the policy, out of the total number of thoughts.

Discussion

The results of Study 3 support the causal claim that social norm information affects policy support by changing query order. This supports Hypothesis 4. As predicted, social norm information affects policy support when queried order is aligned with the condition (i.e., when support thoughts cluster earlier for the control condition and when support thoughts cluster later for the oppose norm treatment condition) and the effect is mitigated when query order is reversed. Query order was established by instruction to be +1 or -1; the effect of this query order

manipulation on query content support the query theory prediction that query order affects query content through retrieval-induced forgetting.

Study 4: Effect of social norm information interventions on real policy decisions in naturalistic information environment

Method

To establish ecological validity and generalize the findings to a new policy domain, Study 4 used a naturalistic information source, television news, to communicate the policy context and norm information. I sought to replicate the query theory mediation of the effects of norm information of Studies 1 and 2 by examining the effect of norm information (support norm versus oppose norm versus a no norm information control) on policy support for large scale solar farms, using the same cognitive process tracing measures. The norm information stimuli were video clips from actual television news segments about large scale solar farms. The main policy support question was, “do you support or oppose a new state policy that would increase the number of large-scale solar farms in your area by lowering taxes for those companies?”. This was answered on the same 5-point Likert scales used in prior studies, from “strongly support” to “strongly oppose”. Videos stimuli can be viewed on OSF²⁵. This study was preregistered (AsPredicted #204826).

Participants

To detect a small effect size ($f = 0.13$) at 80% power as preregistered (AsPredicted #204826), our target was to recruit 540 participants. We collected our sample on Prolific Academic in December 2024 by advertising a survey experiment to respondents located in the United States.

²⁵ https://osf.io/zdmj2/?view_only=d885d9ff45364a5991514b8bb75c804d

We excluded participants who failed one or more of the four attention checks and participants who spent more than 60 minutes on the survey. Of the 586 responses that were collected, we excluded 19. After these preregistered exclusions, we achieved 567 US-based participants (50% female, 75% white, 53% four-year college degree or more; mean age group = 45-55 years, median income group = \$50,000-74,999, 30% Democrats, 38% Independents, 30% Republicans).

Procedure

Study 4 followed the same experimental procedure as Study 1, the only differences were the policy scenario, the fact that the policy and norm stimuli were provided as video information rather than text, and willingness-to-pay questions were not included. There were two additional questions. The first was a manipulation check for the provided norm information, with the following question asked after respondents had answered the policy-support dependent measures: “What proportion of people in the United States do you think support this policy that increases large scale solar farms?,” answered on a scale from 0 to 100. The second was a question about general skepticism: “Overall, to what extent were you skeptical about the accuracy of the information you saw in the video clips?”, on a 4-point Likert scale from “not at all skeptical” to “very skeptical”. This took the place of the question about skepticism toward the norm information in Studies 1 and 2.

Video stimuli were developed by editing and combining television news clips about large-scale solar energy. Video clips were sourced from the Television News Analyzer, hosted by the Internet Archive (Hong et al., 2021). All participants (in control and two treatment conditions) were shown a 51-second clip introducing the policy topic. Next, participants in the support norm condition were shown a clip that is 1 minute and 15 seconds long that communicated social norm information in support of large-scale solar farms; participants in the oppose norm condition were

shown a clip that is 1 minute and 17 seconds long that communicated the social norm information that opposed large-scale solar farms. The abbreviated support norm transcript reads: “*Clip 1*: Solar is the fastest growing renewable energy source, up nearly eight-fold in the last decade.... In the two years since the new rebates, Georgia Interfaith Power and Light says it has seen a 50% increase in faith communities enrolling in their solar-wise program. *Clip 2*: Solar is one of the most popular things in America right now. You know, 89% of Americans support solar. That's about 85% of Republicans, 95% of Democrats.... *Clip 3*: ... according to the Yale program on Climate Change Communication at George Mason University, 78% of registered voters support providing tax rebates for folks who produce energy efficient vehicles or solar panels...” The abbreviated support norm transcript reads: “*Clip 1*: A new report from Pew Research Center finds that although many Americans do want to reduce their carbon footprint, they're actually taking fewer steps to do so. The main part of this study focused on renewable energy, specifically solar and wind. ...*Clip 2*: Well, a brand-new NBC News poll finds that only 5% of voters out of 100 think climate change is the most important issue facing this country right now. *Clip 3*: [a farmer] owns 40 acres of property in southwest Johnson County, Kansas. Now he's facing a massive solar field about 150 feet from his property line, taking away from its rural beauty.... One project in the works in Kansas would be massive, and residents are looking at two years of construction in a populated area and plenty of noise and dust. *Clip 4*: Fewer people are buying rooftop solar panels for their homes. The industry is shrinking, hit by high interest rates, tight credit and cuts to solar incentives. More than 100 solar contractors have already gone out of business during the past year, as demand has dried up.” The full video transcripts are available in Table A4.9 and the videos are available on OSF.

These stimuli were validated in a separate, within-subjects study ($n = 20$) conducted on Prolific Academic²⁶. In this study, the shared policy context video was rated as “neutral” toward solar energy ($M = 3.6$ on a 1 to 7 Likert-scale where 4 = “neutral”), the support norm video was rated as “positive” toward solar energy ($M = 6.3$), and the oppose norm video was rated as “negative” ($M=2.4$). Using mixed-effects linear regression analysis with fixed effects for the stimuli and random effects for participant, we confirmed that, compared to the policy context video ($\beta= 3.60$, $SE = 0.28$, $p < 0.001$), participants perceived that the video’s sentiment toward solar energy was significantly higher in the positive norm video ($\beta= 2.65$, $SE = 0.36$, $p < 0.001$) and significantly lower in the negative norm video ($\beta= -1.18$, $SE = 0.36$, $p = 0.002$) (full regression table in Table A4.10).

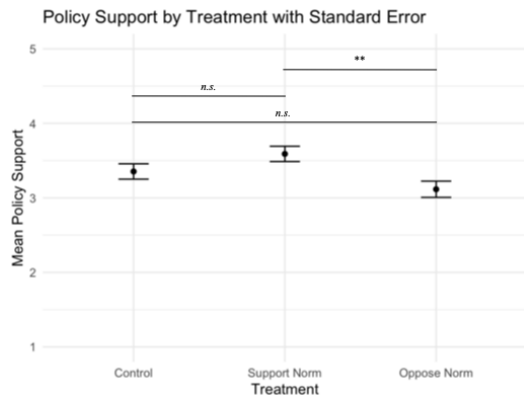
Results

First, I evaluated the manipulation check and found that the stimuli appropriately manipulated norm perceptions of the percentage of Americans supporting the proposed large scale solar policy. The average norm perception in the control group was 53.6%, indicating that participants who saw only the policy context video thought that about half of Americans support the described large scale solar policy. Using linear regression analysis, I found that, compared to the no information control ($\beta= 53.62$, $SE = 1.32$, $p < 0.001$), the positive norm condition increased the perceived norm to 69.4% ($\beta= 15.79$, $SE = 1.87$, $p < 0.001$) and the negative norm condition decreased the perceived norm to 41.5% ($\beta= -12.14$, $SE = 1.87$, $p < 0.001$). Notably, the magnitude of the effect of the two video stimuli on perceived norms was similar.

²⁶ Participants in this stimulus validation study watched each of the three stimuli videos (policy context, support norm information, and oppose norm information) in a randomized order and after each video, answered, “What is the video's sentiment toward solar energy?”, on a 7-point Likert scale from “very negative toward solar energy” to “very positive toward solar energy”.

Replicating results from Studies 1, (positive versus negative) social norm information affected policy support ($F(2,564) = 5.11, p = 0.0063$), supporting Hypothesis 1. Figure 4.5a shows that the supporting norm information treatment increased policy support compared to the opposing norm information treatment ($t(375) = 3.17, p = 0.0016$). The number of thoughts listed was not significantly different between the two norm conditions ($F(1,376) = 0.023, p = .88$) and participants did list fewer thoughts in the control condition than in the two treatment conditions ($M_{\text{support norm}} = 4.81, M_{\text{oppose norm}} = 4.85, M_{\text{control}} = 4.18; F(2,564) = 6.59, p = 0.0015$).

As predicted by query theory and Hypothesis 2, we found that norm information affected query order and balance of support. Query order, measured as SMRD, was significantly higher in the support norm condition ($M_{\text{support norm}} = 0.33, SD_{\text{support norm}} = 0.88$) than in the oppose norm condition ($M_{\text{oppose norm}} = 0.12, SD_{\text{oppose norm}} = 0.94; t(375) = 2.27, p = 0.024$). Higher SMRD indicates that support thoughts clustered earlier in the thought list in the support norm condition than in the oppose norm condition. Balance of policy support, measured as the proportion of support-thoughts, was also significantly higher in the support norm condition ($M_{\text{support norm}} = 0.54, SD_{\text{support norm}} = 0.39$) than in the oppose norm condition ($M_{\text{oppose norm}} = 0.43, SD_{\text{oppose norm}} = 0.40; t(376) = 2.53, p = 0.012$).



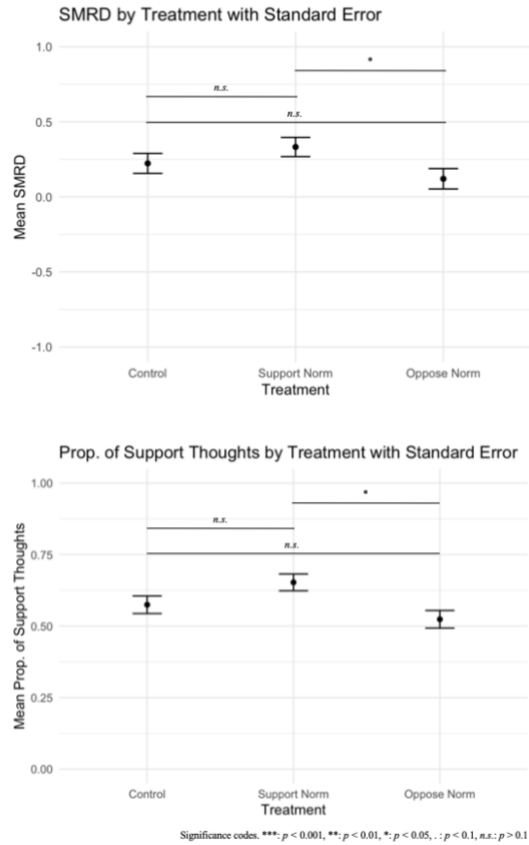


Figure 4.5. Results of Study 4: policy support (top panel), standardized median rank difference (SMRD; middle panel), and proportion of support thoughts (bottom panel) are reported by condition. Policy support is measured on a 5-point Likert scale where 5= strongly support, 3 = neither support nor oppose, and 1= strongly oppose. SMRD is measured on a scale from -1 to 1 where positive values indicate support-thoughts clustered earlier than oppose-thoughts, negative values indicate the opposite, and zero indicates randomly dispersed thoughts. Proportion of support thoughts is measured as the proportion of thoughts listed that are in support of the policy, out of the total number of thoughts.

Skepticism about the policy information moderated the effect of norm information on policy support. This was pre-registered as an exploratory analysis for the present study. Using

linear regression, we regress policy support on treatment (support versus oppose norm information) interacted with level of skepticism. The interaction was significant ($\beta = 0.94$, $SE = 0.15$, $p < 0.0001$), such that higher skepticism attenuated some of the effect of norm treatment on policy support (Figure A4.11). Neither query metric varied by policy skepticism (SMRD: $F(1,565) = 0.50$, $p = 0.48$; Balance of support: $F(1,565) = 0.98$, $p = 0.32$).

A moderated mediation analysis was conducted to examine Hypothesis 3—whether the query mechanisms explain the effect of norm information on policy support while accounting for skepticism. The analysis showed that the effect of positive norm information, compared to negative norm information, on policy support was mediated by both query order and balance of support²⁷. The two query variables were correlated and non-overlapping in this study ($r = .75$). Therefore, a principal-components factor analysis was conducted to calculate a ‘prominence-of-support-thoughts’ factor, following the methodology of prior query theory studies (i.e., Weber et al., 2007). I ran a moderated mediation analysis to examine whether prominence-of-support-thoughts mediated the effect of norm treatment on policy support. Given the moderating role of skepticism on the effect of norm information on policy support, policy skepticism was included as a moderator of the direct and indirect effects. The results suggested partial mediation; there was a significant indirect effect of norm treatment on the mediator ($\beta = 0.27$, 95%CI [0.07,0.48], $p = 0.0070$). The direct effect was marginally significant when accounting for the mediator ($\beta = 0.18$, 95%CI [-0.03,0.38], $p = 0.093$), suggesting partial mediation. Prominence-of-support-thoughts mediated

²⁷ The moderated mediation analysis was also run individually for the two query metrics. In the SMRD analysis, there was a significant indirect effect ($\beta = 0.18$, 95%CI [0.02,0.35], $p = 0.028$) and the direct effect was still significant after accounting for the mediator ($\beta = 0.28$, 95%CI [0.04,0.52], $p = 0.019$), suggesting partial mediation. SMRD mediated 39% of the effect (Table A4.12). In the balance of support thoughts mediation analysis, there was a significant indirect effect ($\beta = 0.27$, 95%CI [0.05,0.48], $p = 0.013$) and the direct effect was still significant after accounting for the mediator ($\beta = 0.20$, 95%CI [0.002, 0.39], $p = 0.047$), suggesting partial mediation. Proportion of support thoughts mediated 57% of the effect (Table A4.13).

61% of the effect of norm information on policy support (Figure 4.6). To examine the moderating effect, I tested whether the average causal mediation effects and direct effects were significantly different between high and low skepticism participants. Both were significant, such that the indirect effect was significantly higher for non-skeptics ($Diff = 0.91, p < 0.001$) and the direct effect after accounting for the mediator was significantly higher for non-skeptics ($Diff = 0.50, p = 0.018$). Taken together, the mediation analysis support Hypothesis 3 in the naturalistic stimuli context when participants were not skeptical of the policy context and information.

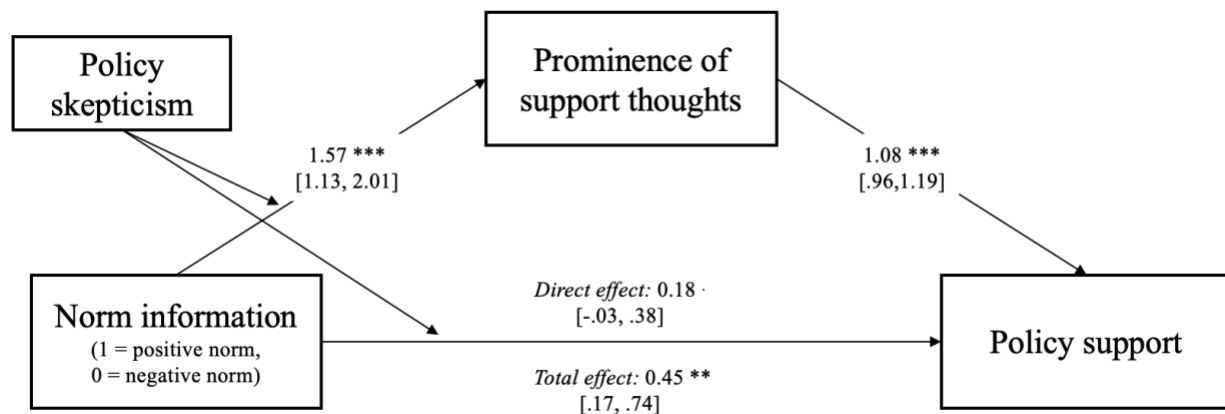


Figure 4.6. Moderated mediation results of Study 4: prominence of support thoughts is investigated as a mediator of and skepticism as a moderator on the effect of treatment group on policy support. The 95% confidence intervals are calculated using 10,000 bootstrap resamples. ACME = 0.27 [.07, .48], $p = .007$; ADE = 0.18 [-0.03, .38], $p = .09$; Total Effect = 0.45 [.17, .74], $p = .002$.

Discussion

Study 4 successfully replicated the conceptual findings from Study 1 in a real policy context using a naturalistic information medium. While Studies 2a-c showed asymmetric effects

that varied by policy topic, Study 4 found symmetric effects of both positive and negative norm information regarding large-scale solar developments. Specifically, norm information supporting large-scale solar policy increased policy support compared to opposing norm information by enhancing the prominence of policy-supporting thoughts for non-skeptics. These consistent results across different contexts provide compelling evidence for the query theory explanation of norm information interventions and highlight the moderating role of skepticism, strengthening the validity and generalizability of our findings.

Study 5: Process paradigm validation

Method

To evaluate whether the thought-listing paradigm of the previous studies captured a naturalistic thought process, a validation study was conducted. In Studies 1-4, individuals' thoughts were captured through thought-listing, where participants type out the thoughts that go through their mind as they make a policy decision. This study investigated whether this paradigm affected the policy support outcome. All participants were exposed to the same policy scenario context and norm information stimuli before being randomly assigned to one of four process conditions: think-only, list thoughts, dictate thoughts, and dictate thoughts with explicit separation, as described below. Study 5 utilized the policy context and negative norm information stimuli about CCS of Study 2c. This study was preregistered (AsPredicted #207593).

Participants

To detect a medium-to-small effect size ($d = 0.18$) at 80% power I targeted a sample size of 400 participants. The sample was collected through Prolific Academic in January 2025 by advertising a survey experiment to respondents located in the United States. Participants were

excluded if they failed one or more of the attention checks ($n = 10$) or if they spent more than 60 minutes on the survey ($n = 5$). 2 participants that were not assigned to any condition due to a technical error by the survey platform were also excluded. The full sample analyzed included 401 participants (52% female, 64% white, 55% four-year college degree or more; mean age group = 45-54 years old, median income group = \$50,000-\$74,999, 33% Democrats, 38% Independents, 27% Republicans).

Procedure

The validation study utilized the policy scenario from Study 2c, and all participants were shown the CCS negative norm information (Table 2A). Participants were randomly assigned to one of four conditions: think-only, list thoughts, dictate thoughts, and dictate thoughts with explicit separation. First, participants received a training on the mechanism used in their condition (i.e., participants in the thought-listing condition practiced the type-aloud method and participants in the two dictation conditions practiced recording themselves with the corresponding instructions). Second, all participants read the CCS policy scenario and were shown the negative social norm information. Third, participants received condition-specific process instructions. In the think-only condition, participants were instructed to “carefully consider the question and policy, and the different arguments for and against the policy”. In the thought-listing condition, participants received the same instructions used in Studies 1-4: “list the thoughts that go through your mind as you evaluate whether you support or oppose the policy above”. In the thought dictation condition, participants received the thought-listing instructions by being asked to record themselves when speaking their thoughts aloud: “now record your audio and begin speaking the thoughts that go through your mind as you evaluate whether you support or oppose the policy above”. In the thought dictation condition with explicit separation, participants received the same instructions as the

dictation condition with the added instruction to “Between each separate thought that you have, please clearly say ‘next’”. Fourth, all participants answered the policy support question (as in Study 2c). And finally, participants answered the demographic and skepticism questions. Participants did not categorize their thoughts.

Results

The results confirmed that the thought-listing paradigm used in the present chapter did not significantly affect policy support compared to the other paradigms. Using a one-way ANOVA, I found that policy support did not vary significantly by process paradigm condition ($F(3, 397) = 0.58, p = 0.63$). Figure A4.14 shows average policy support for each condition. The average policy support across conditions was 3.01 ($SE_{\text{Study 5}} = 0.069$), which was not significantly different from the average policy support in the negative norm group in Study 2c ($M_{\text{Study 2c}} = 2.80, SE_{\text{Study 2c}} = 0.10; t(586) = 1.74, p = 0.083$). Skepticism about the policy context did not vary across conditions ($F(3, 397) = 1.98, p = 0.12$). Skepticism about the norm information was significantly different across conditions ($F(3, 397) = 2.64, p = 0.049$), but post-hoc pairwise comparisons with Bonferroni adjustments reveal that the only significant difference was between the thought listing and thought dictation with explicit separation ($p = 0.049$), with norm skepticism higher in thought listing condition. Finally, the number of thoughts listed in the thought listing and dictate thoughts with explicit separation conditions were compared using a two-sample t-test²⁸. The number of thoughts listed did not vary significantly between the two conditions ($M_{\text{thought listing}} = 4.19, SE_{\text{thought listing}} = 0.21, M_{\text{dictate with separation}} = 4.29, SE_{\text{dictate with separation}} = 0.26; t(136) = 0.28, p = 0.78$).

Discussion

²⁸ These two conditions were the only ones with explicit separation between thoughts, so they are the only two where “number of thoughts” could be calculated.

Study 5 offers evidence that the thought listing paradigm used in Studies 1 through 4 in this chapter captures a naturalistic thought process. The results of Study 5 showed that typing thoughts did not significantly affect policy support compared to verbal thought-listing or no thought listing with only internal reflection. By comparing the number of thoughts listed in the typing and verbal thought listing conditions, this study also suggested that the extent of the deliberation was not affected by typing compared to verbalizing, which might be more natural. Query theory assumes that queries are issued and executed in an automatic and implicit fashion (Johnson et al., 2007; Weber et al., 2007). The thought listing procedure, which is more cumbersome than natural thought, is merely a way to make the implicit explicit. Study 5 shows that this is a plausible assumption in the domain of social cognition.

General Discussion

The five studies reported here examined one cognitive mechanism for the effect of social norm information interventions on policy support. These studies support the extension of query theory into the social cognition domain by examining the premises of query theory (Johnson et al., 2007). Policy decisions were deconstructed into component queries (policy-supporting arguments and policy-opposing arguments) and social norm information interventions were examined for their effect on the order and clustering of those two queries (Study 1; Studies 2a-c; Study 4). Despite the overall popularity of the hypothetical policy in Study 1, thoughts favoring the policy occurred earlier in the positive norm information condition than in the negative norm information condition, where later and less clustered consideration of policy support also reduced the balance of support for the policy. Both query order and balance of thoughts mediated the effect of social norm information on policy support in this hypothetical context. Study 4 extended these findings to a real-world policy context, focusing on large-scale solar farms, with a naturalistic norm

information intervention medium. Across the studies, I found that participants who were skeptical of the policy domain or norm information were not affected by norm information interventions.

Study 3 tested the causal role of query order by manipulating the order of these queries, such that half of the participants in each norm condition first listed policy-supporting thoughts and the other half first listed policy-opposing thoughts. When participants listed thoughts in the ‘natural order’, listing thoughts first that aligned with the norm condition, we replicated the effects of norm information on policy support. When we reversed the query order, i.e., having participants list thoughts first that misaligned with their norm condition, the effect of norm information on policy support was attenuated, with corresponding effects on the balance of support process variable. Study 3 provides evidence for the causal effect of natural query order but also for the fact that policy support is partially malleable by the decision environment, in particular by the way in which it primes the order in which supporting or opposing arguments are solicited.

The results of the real policy contexts in Studies 2a-c, where we analyzed the effect of providing norm information to a no-information control condition, suggests limitations to these norm effects as well as potentially asymmetrical effects. In Studies 2a and 2b, when participants were skeptical of the norm information that was presented, there was no measured effect on policy support or the query metrics. This aligns with the literature which suggests the current levels of partisanship in the United States associated with climate and social policies has reduced trust in science (Azevedo & Jost, 2021). And, that exposure to partisan news sources can decrease trust in the media (Guess et al., 2021). The norm information in these studies was framed as a survey finding communicated through a common media source. These two features, along with the familiarity of the policy domains (i.e., energy efficient appliances and labor organizing protections) may explain this boundary condition. The policy context in Study 2c, carbon capture and storage

technology, is much less familiar to the public (Pianta et al., 2021; Composto et al., *under review*), which might explain why for this study, skeptics like non-skeptics were affected by the norm intervention. Skepticism across the three policy domains was significantly lower in the labor organizing policy condition than in either of the climate-related policies ($X^2(2) = 14.51, p = 0.00071$). This suggests levels of skepticism were related to how politically polarized the policy domain was and the effect of skepticism on norm intervention effectiveness was related to domain familiarity.

I argue that query theory is a process-model to explain social norm information interventions and does not crowd out other norm theories. In particular, the query theory process findings of this chapter offer a process explanation of the focus theory of normative conduct (Cialdini et al., 1991). Social norms give people decisional shortcuts, as proposed in the focus theory of normative conduct, by directing attention, measured as initial queries, to the normative option first, which subsequently affects balance of support. The present studies show the effects to the cognitive process of making the norm salience, through an information intervention. Study 4 exemplifies this process by comparing between positive and negative norm information, suggesting that the norm that is made salient activates memories that are in alignment.

The memory-retrieval theory that I propose is not mutually exclusive with other hypotheses. I investigated whether the social norm information functioned as a typical memory item that appeared in thought lists. To do this, the aspects that were listed in Study 1 and Studies 2a-2c were coded for whether or not a social norm was referenced²⁹. Across the studies that were coded, 17% of participants made a clear reference to social norm information. In Study 1, the

²⁹ This coding was conducted post-hoc by the second author of this chapter (AS) after training and creating a code book.

mentions of norm information did not differ between the support and oppose norm conditions³⁰. Comparing the support norm conditions in Studies 2a and 2b, norm mentions also did not vary by policy topic³¹. The highest rate of norm mentions by far was in the CCS (Study 2c) norm condition—38% of people mentioned the norm information. This suggests that people rely more on, or at least give more attention to, norm information when they hold less certain knowledge, as is the case with policies related to new energy technologies (Collins & Michalski, 1989; Pianta et al., 2021).

Future work should investigate the possible differences between the cognitive processes that underpin descriptive and injunctive norm interventions. As discussed, the theoretical and empirical literature on social norms often examines the motivational mechanisms and intervention effects of these two as distinct. The results here should be interpreted as primarily applicable to descriptive norm interventions; the memory-retrieval process presented here aligns with the accuracy motivation (Jacobson et al., 2011). It is an open question if injunctive norm interventions activate a similar cognitive process and whether social approval or reciprocity may come into play more. Similarly, dynamic norms may have stronger effects on query order because they direct focus to a precise pattern of behavior (Sparkman & Walton, 2017). In summary, this memory-retrieval process theory of social norm interventions can expand and deepen our understanding of how social norms shape collective behaviors.

³⁰ The results of a two-sample test of equality of proportions find a non-significant difference ($\chi^2(1) = 0.63, p = 0.43$).

³¹ $\chi^2(1) = 0.064, p = 0.80$

Dissertation General Discussion

An accurate understanding of the social world is a critical part of societal transitions and coordinated collective action. In the context of the energy transition and threats to democracy, there is evidence of systematic misperceptions that lead to systematic failures in coordination and exacerbate threats to society. This dissertation explored the cognitive processes underlying these social phenomena. The organizing question was, how do people come to understand or form impressions of the social world?

The first aim of this dissertation was to explore the extent to which query theory's memory-retrieval process theory explains social cognition, specifically social norm perceptions. Chapter 1 offered a comprehensive meta-analysis of a process theory for individual-level decision making. By bringing together the existing literature on query theory, I offered strong evidence of the memory-retrieval explanation for frame effects. In short, query theory shows that individual preference is constructed, rather than pre-stored; therefore, decisions or judgments shape the construction process, which, in turn, affects choice. In Chapter 2, I drew on norm theory to argue that a similar memory-retrieval process may inform social cognition. I provided evidence that social norm perceptions, like individual preference, are constructed ad hoc, and I provided preliminary evidence for the memory-retrieval mechanisms. If indeed social norm perceptions are constructed through such processes, this may explain why such misperceptions of others' mental states persist and why these perceptions are affected by one's information environment. This underscores the importance of accurate scientific and policy reporting, particularly when there is expert consensus, as is the case with climate change.

An important distinction between social norm construction and individual preference is that context, or the strength of the existing norm, matters. Thought clustering patterns differed

between normative and counter-normative conditions in Chapter 2. More normative topics showed stronger thought clustering, while counter-normative conditions displayed minimal clustering—a pattern that differed from individual choice contexts. Psychological distance appears to have played a crucial role in norm perception malleability. Norms related to psychologically distant topics (corporate environmental practices) were more malleable and more influenced by framing conditions than psychologically proximate topics (meal choices). Similarly, policy support on familiar topics (Chapter 4), was less affected by social norm information. This suggests that people relied more on other’s opinions when they were less familiar with a topic, i.e., new energy technologies.

The second aim of this dissertation was to leverage the process tracing paradigm to understand the cognitive mechanisms of social norm information. This investigation built on the idea that social perceptions are, in part, constructed ad hoc. Chapter 4 presented evidence that policy decisions are deconstructed into component queries (supporting and opposing arguments), and social norm information affects both the order and content of these queries. When positive norm information was presented, policy-supporting thoughts occurred earlier, whereas negative norm information delayed these supportive considerations, reducing overall policy support. I also found that the query order mechanism was causal. Together, this process theory explanation suggests that social norm information directs attention and consideration to the normative option first during policy consideration. These results offer a cognitive explanation for social and system-level phenomena, such as social tipping. This process explanation also illustrates how false social realities can form—a negative norm about a policy (e.g. that most people oppose large-scale solar farms) changed the evaluation process, even when that norm was not true. Such information can

decrease policy support, building opposition based on inaccurate norm perceptions by changing peoples' implicit cognitive evaluation process.

A critical boundary of my findings in Chapter 4 was that participants who were skeptical of the policy context or norm information were not affected by the intervention. This brings me to the third and final aim of this dissertation—to understand how people process complex and information-dense environments to learn social information. The skepticism finding in Chapter 4 was unsurprising—if someone does not believe or trust information, they are less likely to integrate it into their mental models. Therefore, Chapter 3 turned to the real information sources that people engage with every day—cable news. Cable news is a reasonably reliable source of information about the world to those who watch it regularly (i.e., participants in my media diet sample). I leveraged recent advancements in LLMs and big data text analysis to uncover the amount and type of social information that people are exposed to through the news. The results of Chapter 3 suggest that cable news coverage, especially when discussing democracy, communicates a high volume of social norm information (e.g., that most people do not approve of actions that violate democratic norms). The study I reported in Chapter 3 offered preliminary, correlational evidence that people's mental models of the social world are associated with the norm information communicated on their televisions. I am planning future work to probe the causality of this relationship.

This dissertation presented a process theory for accurate and inaccurate social perceptions. This work highlighted the role of attention and memory-retrieval underlying social cognition of diverse information sources in contexts of complex societal change. In short, people construct mental models of the social world from direct and indirect experiences. When policy topics are less familiar, there is a greater reliance on the evaluation of others and inaccurate norm perceptions can have outsized consequences for societal resilience and transitions. The work I presented in this

dissertation explains how norm information directs attention and memory-retrieval to construct models that align with the presented information.

References

- Abrahamse, W., & Steg, L. (2013). Social influence approaches to encourage resource conservation: A meta-analysis. *Global Environmental Change-Human and Policy Dimensions*, 23(6), 1773–1785. <https://doi.org/10.1016/j.gloenvcha.2013.07.029>
- Adjerid, I., Samat, S., & Acquisti, A. (2016). A Query-Theory Perspective of Privacy Decision Making. *The Journal of Legal Studies*, 45(S2), S97–S121. <https://doi.org/10.1086/689599>
- Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9), 1082–1095.
- Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, 7(3), 522–530. <https://doi.org/10.3758/BF03214366>
- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063–1087.
- Andersson, M., & von Borgstede, C. (2010). Differentiation of determinants of low-cost and high-cost recycling. *Journal of Environmental Psychology*, 30(4), 402–408. <https://doi.org/10.1016/j.jenvp.2010.02.003>
- Andreoni, J., Nikiforakis, N., & Siegenthaler, S. (2021). Predicting social tipping and norm change in controlled experiments. *Proceedings of the National Academy of Sciences*, 118(16), e2014893118. <https://doi.org/10.1073/pnas.2014893118>
- Appelt, K. C., Hardisty, D. J., & Weber, E. U. (2011). Asymmetric discounting of gains and losses: A query theory account. *Journal of Risk and Uncertainty*, 43(2), 107–126.

- Arceneaux, K., & Truex, R. (2023). Donald Trump and the Lie. *Perspectives on Politics*, 21(3), 863–879. <https://doi.org/10.1017/S1537592722000901>
- Avant, L. L., & Helson, H. (1973). Theories of perception. In *Handbook of general psychology* (pp. xv, 1006–xv, 1006). Prentice Hall.
- Azevedo, F., & Jost, J. T. (2021). The ideological basis of antiscientific attitudes: Effects of authoritarianism, conservatism, religiosity, social dominance, and system justification. *Group Processes & Intergroup Relations*, 24(4), 518–549. <https://doi.org/10.1177/1368430221990104>
- Becker, A. (forthcoming). On the Economic Origins of Concerns Over Women’s Chastity. *Review of Economic Studies*. (Pre-published online August 26, 2024.)
- Bergquist, M., Nilsson, A., & Ejelöv, E. (2019). Contest-Based and Norm-Based Interventions: (How) Do They Differ in Attitudes, Norms, and Behaviors? *Sustainability*, 11(2), Article 2. <https://doi.org/10.3390/su11020425>
- Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms* (pp. xvi, 260). Cambridge University Press.
- Binz, M., Alaniz, S., Roskies, A., Aczel, B., Bergstrom, C. T., Allen, C., Schad, D., Wulff, D., West, J. D., Zhang, Q., Shiffrin, R. M., Gershman, S. J., Popov, V., Bender, E. M., Marelli, M., Botvinick, M. M., Akata, Z., & Schulz, E. (2025). How should the advancement of large language models affect the practice of science? *Proceedings of the National Academy of Sciences*, 122(5), e2401227121. <https://doi.org/10.1073/pnas.2401227121>
- Blanton, H., Köblitz, A., & McCaul, K. D. (2008). Misperceptions about Norm Misperceptions: Descriptive, Injunctive, and Affective ‘Social Norming’ Efforts to Change Health Behaviors.

- Social and Personality Psychology Compass*, 2(3), 1379–1399.
<https://doi.org/10.1111/j.1751-9004.2008.00107.x>
- Bohner, G., & Schlüter, L. E. (2014). A Room with a Viewpoint Revisited: Descriptive Norms and Hotel Guests' Towel Reuse Behavior. *PLoS ONE*, 9(8), e104086.
<https://doi.org/10.1371/journal.pone.0104086>
- Bolsen, T., Leeper, T. J., & Shapiro, M. A. (2014). Doing What Others Do: Norms, Science, and Collective Action on Global Warming. *American Politics Research*, 42(1), 65–89.
<https://doi.org/10.1177/1532673X13484173>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (Eds.). (2009). *Introduction to meta-analysis* (Nachdr.). Wiley.
- Bowman, E. (2024, October 20). Hackers steal information from 31 million Internet Archive users. *NPR*. <https://www.npr.org/2024/10/20/nx-s1-5159000/internet-archive-hack-leak-wayback-machine>
- Broadwater, L. (2024, February 6). Dozens of House Republicans Declare Trump's Jan. 6 Actions Were Not Insurrection. *The New York Times*.
<https://www.nytimes.com/2024/02/06/us/politics/house-republicans-trump-jan-6.html>
- Broockman, D. E., & Kalla, J. L. (2024). Selective exposure and echo chambers in partisan television consumption: Evidence from linked viewership, administrative, and survey data. *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12886>
- Burchardi, K. B., & Penczynski, S. P. (2014). Out of your mind: Eliciting individual reasoning in one shot games. *Games and Economic Behavior*, 84, 39–57.
<https://doi.org/10.1016/j.geb.2013.12.005>

- Burchell, K., Rettie, R., & Patel, K. (2013). Marketing social norms: Social marketing and the ‘social norm approach.’ *Journal of Consumer Behaviour*, 12(1), 1–9.
<https://doi.org/10.1002/cb.1395>
- Burgess, M. G., Van Boven, L., Wagner, G., Wong-Parodi, G., Baker, K., Boykoff, M., Converse, B. A., Dilling, L., Gilligan, J. M., Inbar, Y., Markowitz, E., Moyer, J. D., Newton, P., Raimi, K. T., Shrum, T., & Vandenbergh, M. P. (2024). Supply, demand and polarization challenges facing US climate policies. *Nature Climate Change*. <https://doi.org/10.1038/s41558-023-01906-y>
- Bursztyn, L., Rao, A., Roth, C., & Yanagizawa-Drott, D. (2023). Opinions as Facts. *The Review of Economic Studies*, 90(4), 1832–1864. <https://doi.org/10.1093/restud/rdac065>
- Campbell, B. M., Beare, D. J., Bennett, E. M., Hall-Spencer, J. M., Ingram, J. S. I., Jaramillo, F., Ortiz, R., Ramankutty, N., Sayer, J. A., & Shindell, D. (2017). Agriculture production as a major driver of the Earth system exceeding planetary boundaries. *Ecology and Society*, 22(4), art8. <https://doi.org/10.5751/ES-09595-220408>
- Castilla-Rho, J. C., Rojas, R., Andersen, M. S., Holley, C., & Mariethoz, G. (2017). Social tipping points in global groundwater management. *Nature Human Behaviour*, 1(9), 640–649. <https://doi.org/10.1038/s41562-017-0181-7>
- Charpentier, C. J., De Neve, J.-E., Li, X., Roiser, J. P., & Sharot, T. (2016). Models of affective decision making: How do feelings predict choice? *Psychological Science*, 27(6), 763–775.
<https://doi.org/10.1177/0956797616634654>
- Cheung, M. W.-L. (2014). Fixed- and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods*, 46(1), 29–40.
<https://doi.org/10.3758/s13428-013-0361-y>

- Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(1), 591–621.
<https://doi.org/10.1146/annurev.psych.55.090902.142015>
- Cialdini, R. B., & Jacobson, R. P. (2021). Influences of social norms on climate change-related behaviors. *Current Opinion in Behavioral Sciences*, 42, 1–8.
<https://doi.org/10.1016/j.cobeha.2021.01.005>
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. In *Advances in Experimental Social Psychology* (Vol. 24, pp. 201–234). Elsevier.
[https://doi.org/10.1016/S0065-2601\(08\)60330-5](https://doi.org/10.1016/S0065-2601(08)60330-5)
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. <https://doi.org/10.1037/0022-3514.58.6.1015>
- Clark, C. S., Paluck, E. L., Westwood, S. J., Sen, M., Malhotra, N., & Jessee, S. (2023). Effects of a US Supreme Court ruling to restrict abortion rights. *Nature Human Behaviour*, 1–9.
<https://doi.org/10.1038/s41562-023-01708-4>
- Clayton, K., Davis, N. T., Nyhan, B., Porter, E., Ryan, T. J., & Wood, T. J. (2021). Elite rhetoric can undermine democratic norms. *Proceedings of the National Academy of Sciences*, 118(23), e2024125118. <https://doi.org/10.1073/pnas.2024125118>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Collins, A., & Michalski, R. (1989). The logic of plausible reasoning: A core theory. *Cognitive Science*, 13(1), 1–49.

- Cologna, V., Mede, N. G., Berger, S., Besley, J., Brick, C., Joubert, M., Maibach, E. W., Mihelj, S., Oreskes, N., Schäfer, M. S., van der Linden, S., Abdul Aziz, N. I., Abdulsalam, S., Shamsi, N. A., Aczel, B., Adinugroho, I., Alabrese, E., Aldoh, A., Alfano, M., ... Zwaan, R. A. (2025). Trust in scientists and their role in society across 68 countries. *Nature Human Behaviour*, 1–18. <https://doi.org/10.1038/s41562-024-02090-5>
- Cologna, V., & Siegrist, M. (2020). The role of trust for climate change mitigation and adaptation behaviour: A meta-analysis. *Journal of Environmental Psychology*, 69, 101428. <https://doi.org/10.1016/j.jenvp.2020.101428>
- Composto, J. W., Duncan, S., Johnson, E. J., & Weber, E. U. (n.d.). *A Meta-Analysis of Query Theory, a Psychological Process Account of Framing Effects*.
- Composto, J. W., Greig, C., & Weber, E. U. (n.d.). *Timely information about Carbon Capture and Storage can increase willingness to accept the technology by changing perceived risks and benefits*. SSRN. Retrieved February 24, 2025, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4911410
- Constantino, S. M., & Weber, E. U. (2021). Decision-making under the deep uncertainty of climate change: The psychological and political agency of narratives. *Current Opinion in Psychology*, 42, 151–159. <https://doi.org/10.1016/j.copsyc.2021.11.001>
- Costa, D. L., & Kahn, M. E. (2013). Energy conservation “nudges” and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association*, 11(3), 680–702.
- Costa-Gomes, M. A., & Crawford, V. P. (2006). Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *The American Economic Review*, 96(5).

- de Groot, J. I. M., & Schuitema, G. (2012). How to make the unpopular popular? Policy characteristics, social norms and the acceptability of environmental policies. *Environmental Science & Policy*, 19–20, 100–107. <https://doi.org/10.1016/j.envsci.2012.03.004>
- Dempster, F. N. (1995). 1 - Interference and inhibition in cognition: An historical perspective. In F. N. Dempster, C. J. Brainerd, & C. J. Brainerd (Eds.), *Interference and Inhibition in Cognition* (pp. 3–26). Academic Press. <https://doi.org/10.1016/B978-012208930-5/50002-7>
- Demszky, D., Yang, D., Yeager, D. S., Bryan, C. J., Clapper, M., Chandhok, S., Eichstaedt, J. C., Hecht, C., Jamieson, J., Johnson, M., Jones, M., Krettek-Cobb, D., Lai, L., Jones Mitchell, N., Ong, D. C., Dweck, C. S., Gross, J. J., & Pennebaker, J. W. (2023). Using large language models in psychology. *Nature Reviews Psychology*, 2(11), 688–701. <https://doi.org/10.1038/s44159-023-00241-5>
- Dinner, I., Johnson, E. J., Goldstein, D. G., & Liu, K. (2011). Partitioning default effects: Why people choose not to choose. *Journal of Experimental Psychology: Applied*, 17(4), 332–341. <https://doi.org/10.1037/a0024354>
- Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How Elite Partisan Polarization Affects Public Opinion Formation. *American Political Science Review*, 107(1), 57–79. <https://doi.org/10.1017/S0003055412000500>
- Duncan, S., & Weber, E. U. (n.d.). *Employee Uptake of Company Wellness Programs: Defaults Signal Implied Endorsement, Which Influence Wellness Decisions via Query Theory Mechanisms*.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341x.2000.00455.x>

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634.

<https://doi.org/10.1136/bmj.315.7109.629>

Eichstaedt, J. C., Kern, M. L., Yaden, D. B., Schwartz, H. A., Giorgi, S., Park, G., Hagan, C. A., Tobolsky, V. A., Smith, L. K., Buffone, A., Iwry, J., Seligman, M. E. P., & Ungar, L. H. (2021). Closed- and open-vocabulary approaches to text analysis: A review, quantitative comparison, and recommendations. *Psychological Methods*, 26(4), 398–427.

<https://doi.org/10.1037/met0000349>

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Revised edition). The MIT Press.

Executive Order data as of January 20, 2025. (2025). [Dataset]. Federal Register.

[whitehouse.gov](https://www.whitehouse.gov)

Fasching, N., Iyengar, S., Lelkes, Y., & Westwood, S. J. (2024). Persistent polarization: The unexpected durability of political animosity around US elections. *Science Advances*.

<https://doi.org/10.1126/sciadv.adm9198>

Ferraro, P. J., Miranda, J. J., & Price, M. K. (2011). The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment. *American Economic Review*, 101(3), 318–322.

<https://doi.org/10.1257/aer.101.3.318>

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., &

- Druckman, J. N. (2020). Political sectarianism in America. *Science*, 370(6516), 533–536.
<https://doi.org/10.1126/science.abe1715>
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316–344. <https://doi.org/10.1037/a0021663>
- Geiger, N., & Swim, J. K. (2016). Climate of silence: Pluralistic ignorance as a barrier to climate change discussion. *Journal of Environmental Psychology*, 47, 79–90.
<https://doi.org/10.1016/j.jenvp.2016.05.002>
- Gelfand, M. J., Gavrillets, S., & Nunn, N. (2024). Norm Dynamics: Interdisciplinary Perspectives on Social Norm Emergence, Persistence, and Change. *Annual Review of Psychology*, 75(1), annurev-psych-033020-013319. <https://doi.org/10.1146/annurev-psych-033020-013319>
- Gerber, A. S., & Rogers, T. (2009). Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You. *The Journal of Politics*, 71(1), 178–191.
<https://doi.org/10.1017/S0022381608090117>
- Giese, H., & Hoffmann, J. A. (2023). Socializing social sampling models: The limits of explaining norm perceptions and biases with sampling from social circles. *PLOS ONE*, 18(6), e0286304. <https://doi.org/10.1371/journal.pone.0286304>
- Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research*, 35(3), 472–482. <https://doi.org/10.1086/586910>
- Gross, J., & Vostroknutov, A. (2022). Why do people follow social norms? *Current Opinion in Psychology*, 44, 1–6. <https://doi.org/10.1016/j.copsyc.2021.08.016>

- Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021). The consequences of online partisan media. *Proceedings of the National Academy of Sciences*, 118(14), e2013464118.
<https://doi.org/10.1073/pnas.2013464118>
- Hall, M. E. K., & Druckman, J. N. (2023). Norm-violating rhetoric undermines support for participatory inclusiveness and political equality among Trump supporters. *Proceedings of the National Academy of Sciences*, 120(40), e2311005120.
<https://doi.org/10.1073/pnas.2311005120>
- Hardisty, D. J., Appelt, K. C., & Weber, E. U. (2013). Good or Bad, We Want it Now: Fixed-cost Present Bias for Gains *and* Losses Explains Magnitude Asymmetries in Intertemporal Choice: Good or Bad, We Want it Now. *Journal of Behavioral Decision Making*, 26(4), 348–361. <https://doi.org/10.1002/bdm.1771>
- Hassell, H. J. G., Miles, M. R., & Reuning, K. (2022). Does the Ideology of the Newsroom Affect the Provision of Media Slant? *Political Communication*.
<https://www.tandfonline.com/doi/full/10.1080/10584609.2021.1986613>
- Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In *Social psychology: Handbook of basic principles* (pp. 133–168). The Guilford Press.
- Higgins, E. T., & Lurie, L. (1983). Context, categorization, and recall: The “change-of-standard” effect. *Cognitive Psychology*, 15(4), 525–547. [https://doi.org/10.1016/0010-0285\(83\)90018-X](https://doi.org/10.1016/0010-0285(83)90018-X)
- Higgins, J., Li, T., & Deeks, J. (2024). Chapter 6: Choosing effect measures and computing estimates of effect. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, &

- V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (version 6.5).
Cochrane. www.training.cochrane.org/handbook.
- Higgins, J., & Thomas, J. (2023). Chapter 10: Analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions* (6.4).
<https://training.cochrane.org/handbook/current/chapter-10>
- Higgs, S. (2016). Cognitive processing of food rewards. *Appetite*, 104, 10–17.
<https://doi.org/10.1016/j.appet.2015.10.003>
- Highhouse, S. (1994). A verbal protocol analysis of choice under ambiguity. *Journal of Economic Psychology*, 15(4), 621–635. [https://doi.org/10.1016/0167-4870\(94\)90014-0](https://doi.org/10.1016/0167-4870(94)90014-0)
- Holliday, D. E., Iyengar, S., Lelkes, Y., & Westwood, S. J. (2024). Uncommon and nonpartisan: Antidemocratic attitudes in the American public. *Proceedings of the National Academy of Sciences*, 121(13), e2313013121. <https://doi.org/10.1073/pnas.2313013121>
- Hong, J., Crichton, W., Zhang, H., Fu, D. Y., Ritchie, J., Barenholtz, J., Hannel, B., Yao, X., Murray, M., Moriba, G., Agrawala, M., & Fatahalian, K. (2021). Analysis of Faces in a Decade of US Cable TV News. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 3011–3021.
<https://doi.org/10.1145/3447548.3467134>
- H.R.3684 - 117th Congress (2021-2022): Infrastructure Investment and Jobs Act (2021).
<https://www.congress.gov/bill/117th-congress/house-bill/3684>
- H.R.5376 - 117th Congress (2021-2022): Inflation Reduction Act of 2022 (2022).
<https://www.congress.gov/bill/117th-congress/house-bill/5376/text>

- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods*, 11(2), 193–206. <https://doi.org/10.1037/1082-989X.11.2.193>
- Hunger, S., & Paxton, F. (2022). What's in a buzzword? A systematic review of the state of populism research in political science. *Political Science Research and Methods*, 10(3), 617–633. <https://doi.org/10.1017/psrm.2021.44>
- Irish, M., & Piguet, O. (2013). The Pivotal Role of Semantic Memory in Remembering the Past and Imagining the Future. *Frontiers in Behavioral Neuroscience*, 7. <https://doi.org/10.3389/fnbeh.2013.00027>
- Ivanova, D., Barrett, J., Wiedenhofer, D., Macura, B., Callaghan, M., & Creutzig, F. (2020). Quantifying the potential for climate change mitigation of consumption options. *Environmental Research Letters*, 15(9), 093001. <https://doi.org/10.1088/1748-9326/ab8589>
- Iyengar, S., Leikes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, 22(1), 129–146. <https://doi.org/10.1146/annurev-polisci-051117-073034>
- Jacobson, R. P., Mortensen, C. R., & Cialdini, R. B. (2011). Bodies obliged and unbound: Differentiated response tendencies for injunctive and descriptive social norms. *Journal of Personality and Social Psychology*, 100(3), 433–448. <https://doi.org/10.1037/a0021470>
- Johnson, E. J., & Goldstein, D. (2003). Do Defaults Save Lives? *Science*, 302(5649), 1338–1339. <https://doi.org/10.1126/science.1091721>
- Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 461–474. <https://doi.org/10.1037/0278-7393.33.3.461>

- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136–153. <https://doi.org/10.1037/0033-295X.93.2.136>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263. <https://doi.org/10.2307/1914185>
- Kahneman, D., & Tversky, A. (2000). *Choices, Values, and Frames*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803475>
- Kalkstein, D. A., Hook, C. J., Hard, B. M., & Walton, G. M. (2023). Social norms govern what behaviors come to mind—And what do not. *Journal of Personality and Social Psychology*, 124(6), 1203–1229. <https://doi.org/10.1037/pspi0000412>
- Kapur, S., Wong, S., Leach, B., & Edelman, A. (2025, January 28). Republicans largely back Trump’s attempted federal aid freeze—With some notes of caution. *NBC News*. <https://www.nbcnews.com/politics/congress/republicans-largely-back-trump-attempted-federal-aid-freeze-rcna189682>
- Keizer, K., & Schultz, P. W. (2018). Social Norms and Pro-Environmental Behaviour. In L. Steg & J. I. M. Groot (Eds.), *Environmental Psychology* (1st ed., pp. 179–188). Wiley. <https://doi.org/10.1002/9781119241072.ch18>
- Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, 22(7), 651–665. <https://doi.org/10.1016/j.tics.2018.04.003>
- Kim, E., Lelkes, Y., & McCrain, J. (2022). Measuring dynamic media bias. *Proceedings of the National Academy of Sciences*, 119(32), e2202197119. <https://doi.org/10.1073/pnas.2202197119>

- King, G., Schneer, B., & White, A. (2017). How the news media activate public expression and influence national agendas. *Science*, 358(6364), 776–780.
<https://doi.org/10.1126/science.aao1100>
- Knoll, M. A. Z., Appelt, K. C., Johnson, E. J., & Westfall, J. E. (2015). Time to retire: Why Americans claim benefits early & how to encourage delay. *Behavioral Science & Policy*, 1(1), 53–62. <https://doi.org/10.1177/237946151500100108>
- Lapinski, M. K., & Rimal, R. N. (2005). An Explication of Social Norms. *Communication Theory*, 15(2), 127–147. <https://doi.org/10.1111/j.1468-2885.2005.tb00329.x>
- Lazarsfeld, P., & Merton, R. (1948). Mass communication, popular taste and organized social action. In L. Bryson (Ed.), *The communication of ideas* (pp. 95–118). Harper.
<https://www.semanticscholar.org/paper/Mass-communication-popular-taste-and-organized-Lazarsfeld-Merton/50c6e01e5e345853eab8bebec12b5d95b398f16a>
- Ledgerwood, A., & Callahan, S. P. (2012). The Social Side of Abstraction: Psychological Distance Enhances Conformity to Group Norms. *Psychological Science*, 23(8), 907–913.
<https://doi.org/10.1177/0956797611435920>
- Lees, J., & Cikara, M. (2020). Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nature Human Behaviour*, 4(3), 279–286.
<https://doi.org/10.1038/s41562-019-0766-4>
- Legros, S., & Cislighi, B. (2020). Mapping the Social-Norms Literature: An Overview of Reviews. *Perspectives on Psychological Science*, 15(1), 62–80.
<https://doi.org/10.1177/1745691619866455>
- Li, Y., Gao, J., Enkavi, A. Z., Zaval, L., Weber, E. U., & Johnson, E. J. (2015). Sound credit scores and financial decisions despite cognitive aging. *PNAS Proceedings of the National*

- Academy of Sciences of the United States of America*, 112(1), 65–69.
<https://doi.org/10.1073/pnas.1413570112>
- Lipsey, M. W., & Wilson, D. B. (2009). *Practical meta-analysis* (Nachdr.). SAGE Publ.
- Littvay, L., McCoy, J. L., & Simonovits, G. (2024). It's Not Just Trump: Americans of Both Parties Support Liberal Democratic Norm Violations More Under Their Own President. *Public Opinion Quarterly*, 88(3), 1044–1058. <https://doi.org/10.1093/poq/nfae042>
- Loewenstein, G. F. (1988). Frames of Mind in Intertemporal Choice. *Management Science*, 34(2), 200–214. <https://doi.org/10.1287/mnsc.34.2.200>
- Lüdtke, D. (2016). *esc: Effect Size Computation for Meta Analysis* (p. 0.5.1) [Dataset].
<https://doi.org/10.32614/CRAN.package.esc>
- Lynn, S. K., & Barrett, L. F. (2014). “Utilizing” Signal Detection Theory. *Psychological Science*, 25(9), 1663–1673. <https://doi.org/10.1177/0956797614541991>
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, 58, 593. <https://doi.org/10.1146/annurev.psych.58.110405.085542>
- MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30(1), 41–62.
https://doi.org/10.1207/s15327906mbr3001_3
- Majd, S., Conley, M. A., Weber, E. U., & Andlinger, G. R. (2017). Evidence of Query Theory as a Tool to Assist Restrained Eaters. *American Journal of Health Behavior*, 41(1), 33–41.
<https://doi.org/10.5993/AJHB.41.1.3>
- McGowan, J. A. (2022). The Trouble with Tibble: Environmental, Social, and Governance (ESG) and Fiduciary Duty. *University of Chicago Law School*.

- McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review*, 10(3), 596–602. <https://doi.org/10.3758/bf03196520>
- Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, 126(6), 910–924. <https://doi.org/10.1037/0033-2909.126.6.910>
- Mildenberger, M., & Tingley, D. (2019). Beliefs about Climate Beliefs: The Importance of Second-Order Opinions for Climate Politics. *British Journal of Political Science*, 49(4), 1279–1307. <https://doi.org/10.1017/S0007123417000321>
- Miller, D. T., & Prentice, D. A. (2016). Changing Norms to Change Behavior. *Annual Review of Psychology*, 67(1), 339–361. <https://doi.org/10.1146/annurev-psych-010814-015013>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269, W64. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>
- Mortensen, C. R., Neel, R., Cialdini, R. B., Jaeger, C. M., Jacobson, R. P., & Ringel, M. M. (2019). Trending Norms: A Lever for Encouraging Behaviors Performed by the Minority. *Social Psychological and Personality Science*, 10(2), 201–210. <https://doi.org/10.1177/1948550617734615>
- Muise, D., Hosseinmardi, H., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2022). Quantifying partisan news diets in Web and TV audiences. *Science Advances*, 8(28), eabn0083. <https://doi.org/10.1126/sciadv.abn0083>
- Mundt, D., Batzke, M. C. L., Bläsing, T. M., Gomera Deaño, S., & Helfers, A. (2024). Effectiveness and context dependency of social norm interventions: Five field experiments

- on nudging pro-environmental and pro-social behavior. *Frontiers in Psychology*, 15.
<https://doi.org/10.3389/fpsyg.2024.1392296>
- News Platform Fact Sheet. (2024, September 17). *Pew Research Center*.
<https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/>
- Nielsen, K. S., Nicholas, K. A., Creutzig, F., Dietz, T., & Stern, P. C. (2021). The role of high-socioeconomic-status people in locking in or rapidly reducing energy-driven greenhouse gas emissions. *Nature Energy*, 1–6. <https://doi.org/10.1038/s41560-021-00900-y>
- Nolan, J. M. (2021). Social norm interventions as a tool for pro-climate change. *Current Opinion in Psychology*, 42, 120–125. <https://doi.org/10.1016/j.copsyc.2021.06.001>
- Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W. N., Arrow, K. J., Barrett, S., Carpenter, S., Chapin, F. S., Crépin, A.-S., Daily, G., Ehrlich, P., Folke, C., Jager, W., Kautsky, N., Levin, S. A., Madsen, O. J., ... De Zeeuw, A. (2016). Social norms as solutions. *Science*, 354(6308), 42–43.
<https://doi.org/10.1126/science.aaf8317>
- Nyborg, K., & Rege, M. (2003). On social norms: The evolution of considerate smoking behavior. *Journal of Economic Behavior & Organization*, 52(3), 323–340.
[https://doi.org/10.1016/S0167-2681\(03\)00031-3](https://doi.org/10.1016/S0167-2681(03)00031-3)
- Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, 14(3).
- Otto, I. M., Donges, J. F., Cremades, R., Bhowmik, A., Hewitt, R. J., Lucht, W., Rockström, J., Allerberger, F., McCaffrey, M., Doe, S. S. P., Lenferna, A., Morán, N., van Vuuren, D. P., & Schellnhuber, H. J. (2020). Social tipping dynamics for stabilizing Earth's climate by 2050.

- Proceedings of the National Academy of Sciences*, 117(5), 2354–2365.
<https://doi.org/10.1073/pnas.1900577117>
- Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3), 574–587.
<https://doi.org/10.1037/a0011989>
- Paluck, E. L., & Shepherd, H. (2012). The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology*, 103(6), 899–915. <https://doi.org/10.1037/a0030015>
- Paluck, E. L., Shepherd, H., & Aronow, P. M. (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proceedings of the National Academy of Sciences*, 113(3), 566–571. <https://doi.org/10.1073/pnas.1514483113>
- Pasek, M. H., Ankori-Karlinsky, L.-O., Levy-Vene, A., & Moore-Berg, S. L. (2022). Misperceptions about out-partisans’ democratic values may erode democracy. *Scientific Reports*, 12(1), 16284. <https://doi.org/10.1038/s41598-022-19616-4>
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker* (pp. xiii, 330). Cambridge University Press. <https://doi.org/10.1017/CBO9781139173933>
- Pianta, S., Rinscheid, A., & Weber, E. U. (2021). Carbon Capture and Storage in the United States: Perceptions, preferences, and lessons for policy. *Energy Policy*, 151, 112149. <https://doi.org/10.1016/j.enpol.2021.112149>
- Poon, N., Luckman, A., Isoni, A., & Mullett, T. L. (2023). A query theory account of the attraction effect. *Cognition*, 238, 105495. <https://doi.org/10.1016/j.cognition.2023.105495>

- Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2), 243–256. <https://doi.org/10.1037/0022-3514.64.2.243>
- Pryor, C., Perfors, A., & Howe, P. D. L. (2019). Even arbitrary norms influence moral decision-making. *Nature Human Behaviour*, 3(1), 57–62. <https://doi.org/10.1038/s41562-018-0489-y>
- Rathje, S., Mirea, D.-M., Sucholutsky, I., Marjeh, R., Robertson, C. E., & Van Bavel, J. J. (2024). GPT is an effective tool for multilingual psychological text analysis. *Proceedings of the National Academy of Sciences*, 121(34), e2308950121. <https://doi.org/10.1073/pnas.2308950121>
- Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From Knowing to Remembering: The Semantic–Episodic Distinction. *Trends in Cognitive Sciences*, 23(12), 1041–1057. <https://doi.org/10.1016/j.tics.2019.09.008>
- Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning: Current research and theory*. Appleton-Century-Crofts. <https://www.semanticscholar.org/paper/A-theory-of-Pavlovian-conditioning-%3A-Variations-in-Rescorla/f37a4126cdeeda8e36b80612578fb7ef46f9ce94>
- Rhodes, N., Shulman, H. C., & McClaran, N. (2020). Changing Norms: A Meta-Analytic Integration of Research on Social Norms Appeals. *Human Communication Research*, 46(2–3), 161–191. <https://doi.org/10.1093/hcr/hqz023>
- Rips, L. J., & Turnbull, W. (1980). How big is big? Relative and absolute properties in memory. *Cognition*, 8(2), 145–174. [https://doi.org/10.1016/0010-0277\(80\)90010-4](https://doi.org/10.1016/0010-0277(80)90010-4)

- Ross, L., & Anderson, C. A. (1982). Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 129–152). Cambridge University Press.
- Roth, P. L., Le, H., Oh, I.-S., Van Iddekinge, C. H., & Bobko, P. (2018). Using beta coefficients to impute missing correlations in meta-analysis research: Reasons for caution. *Journal of Applied Psychology*, 103(6), 644–658. <https://doi.org/10.1037/apl0000293>
- Rothstein, H., Sutton, A., & Borenstein, M. (2006). Publication Bias in Meta-analysis. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169(4), 1012–1012. https://doi.org/10.1111/j.1467-985X.2006.00446_14.x
- Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & Cognition*, 17(6), 759–769. <https://doi.org/10.3758/BF03202637>
- Schulte-Mecklenbeck, M., Johnson, J. G., Böckenholt, U., Goldstein, D. G., Russo, J. E., Sullivan, N. J., & Willemsen, M. C. (2017). Process-Tracing Methods in Decision Making: On Growing Up in the 70s. *Current Directions in Psychological Science*, 26(5), 442–450. <https://doi.org/10.1177/0963721417708229>
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5), 429–434. <https://doi.org/10.1111/j.1467-9280.2007.01917.x>
- Sherif, M. (1937). An experimental approach to the study of attitudes. *Sociometry*, 1, 90–98.
- Shrout, P., & Bolger, N. (2002). Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations. *Psychological Methods*, 7, 422–445. <https://doi.org/10.1037/1082-989X.7.4.422>

- Shteynberg, G., Bramlett, J. M., Fles, E. H., & Cameron, J. (2016). The broadcast of shared attention and its impact on political persuasion. *Journal of Personality and Social Psychology, 111*(5), 665–673. <https://doi.org/10.1037/pspa0000065>
- Smith, J., Louis, W., Terry, D., Greenaway, K., Clarke, M., & Cheng, X. (2012). Congruent or conflicted? The impact of injunctive and descriptive norms on environmental intentions. *Journal of Environmental Psychology, 32*(4), 353–361.
- Smith, J. R., & Louis, W. R. (2008). Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude–behaviour relationship. *British Journal of Social Psychology, 47*(4), 647–666. <https://doi.org/10.1348/014466607X269748>
- Sparkman, G., Geiger, N., & Weber, E. U. (2022). Americans experience a false social reality by underestimating popular climate policy support by nearly half. *Nature Communications, 13*(1), Article 1. <https://doi.org/10.1038/s41467-022-32412-y>
- Sparkman, G., & Walton, G. M. (2017). Dynamic Norms Promote Sustainable Behavior, Even if It Is Counternormative. *Psychological Science, 28*(11), 1663–1674. <https://doi.org/10.1177/0956797617719950>
- Sparkman, G., & Walton, G. M. (2019). Witnessing change: Dynamic norms help resolve diverse barriers to personal change. *Journal of Experimental Social Psychology, 82*, 238–252. <https://doi.org/10.1016/j.jesp.2019.01.007>
- Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87*(1), 96–111. <https://doi.org/10.1037/0021-9010.87.1.96>
- Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye Movements in Risky Choice. *Journal of Behavioral Decision Making, 29*(2–3), 116–136. <https://doi.org/10.1002/bdm.1854>

- Tankard, M. E., & Paluck, E. L. (2016). Norm Perception as a Vehicle for Social Change: Vehicle for Social Change. *Social Issues and Policy Review*, 10(1), 181–211.
<https://doi.org/10.1111/sipr.12022>
- Tankard, M. E., & Paluck, E. L. (2017). The Effect of a Supreme Court Decision Regarding Gay Marriage on Social Norms and Personal Attitudes. *Psychological Science*, 28(9), 1334–1344. <https://doi.org/10.1177/0956797617709594>
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>
- 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.
- Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21(11), 1559–1573.
<https://doi.org/10.1002/sim.1187>
- Ting, H., & Wallsten, T. S. (2011). A query theory account of the effect of memory retrieval on the sunk cost bias. *Psychonomic Bulletin & Review*, 18(4), 767–773.
<https://doi.org/10.3758/s13423-011-0099-4>
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, 59, 1–38.
<https://doi.org/10.18637/jss.v059.i05>
- Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory*, 1, 381–403.
- Tulving, E. (1983). *Elements of Episodic Memory*. Oxford University Press.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*, 40(4), 385.

- Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, 53(1), 1–25. <https://doi.org/10.1146/annurev.psych.53.100901.135114>
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
<https://doi.org/10.1007/BF00122574>
- Van Boven, L., Ehret, P. J., & Sherman, D. K. (2018). Psychological Barriers to Bipartisan Public Support for Climate Policy. *Perspectives on Psychological Science*, 13(4), 492–507.
<https://doi.org/10.1177/1745691617748966>
- Vasconcelos, V. V., Constantino, S. M., Dannenberg, A., Lumkowsky, M., Weber, E., & Levin, S. (2021). Segregation and clustering of preferences erode socially beneficial coordination. *Proceedings of the National Academy of Sciences*, 118(50), e2102153118.
<https://doi.org/10.1073/pnas.2102153118>
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36, 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Voelkel, J. G., Stagnaro, M. N., Chu, J. Y., Pink, S. L., Mernyk, J. S., Redekopp, C., Ghezae, I., Cashman, M., Adjodah, D., Allen, L. G., Allis, L. V., Baleria, G., Ballantyne, N., Van Bavel, J. J., Blunden, H., Braley, A., Bryan, C. J., Celniker, J. B., Cikara, M., ... Willer, R. (2024). Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science*, 386(6719), eadh4764. <https://doi.org/10.1126/science.adh4764>
- Von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior* (Third). Princeton University Press.

- Walters, D. J., & Hershfield, H. E. (2020). Consumers Make Different Inferences and Choices When Product Uncertainty Is Attributed to Forgetting Rather than Ignorance. *Journal of Consumer Research*, 47(1), 56–78. <https://doi.org/10.1093/jcr/ucz053>
- Wang, J. T., Spezio, M., & Camerer, C. F. (2010). Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games. *American Economic Review*, 100(3), 984–1007. <https://doi.org/10.1257/aer.100.3.984>
- Weber, E. U., & Johnson, E. J. (2006). Constructing Preferences from Memory. In S. Lichtenstein & P. Slovic (Eds.), *The Construction of Preference* (pp. 397–410). Cambridge University Press. <https://papers.ssrn.com/abstract=1301075>
- Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric Discounting in Intertemporal Choice: A Query-Theory Account. *Psychological Science*, 18(6), 516–523. <https://doi.org/10.1111/j.1467-9280.2007.01932.x>
- Westhoek, H., Lesschen, J. P., Rood, T., Wagner, S., De Marco, A., Murphy-Bokern, D., Leip, A., van Grinsven, H., Sutton, M. A., & Oenema, O. (2014). Food choices, health and environment: Effects of cutting Europe's meat and dairy intake. *Global Environmental Change*, 26, 196–205. <https://doi.org/10.1016/j.gloenvcha.2014.02.004>
- Westwood, S. J., & Lelkes, Y. (2024). *America's Political Pulse* [Dataset]. <https://polarizationresearchlab.org/americas-political-pulse/>
- Widmann, T., & Wich, M. (2023). Creating and Comparing Dictionary, Word Embedding, and Transformer-Based Models to Measure Discrete Emotions in German Political Text. *Political Analysis*, 31(4), 626–641. <https://doi.org/10.1017/pan.2022.15>
- Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J., Hawkes,

- C., Zurayk, R., Rivera, J. A., Vries, W. D., Sibanda, L. M., ... Murray, C. J. L. (2019). Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet*, 393(10170), 447–492. [https://doi.org/10.1016/S0140-6736\(18\)31788-4](https://doi.org/10.1016/S0140-6736(18)31788-4)
- Wilson, D. (2023). *Practical Meta Analysis Effect Size Calculator* (Version 2023.11.27) [Computer software].
- Wolfswinkel, S., Raghoobar, S., Dagevos, H., De Vet, E., & Poelman, M. P. (2024). How perceptions of meat consumption norms differ across contexts and meat consumer groups. *Appetite*, 195, 107227. <https://doi.org/10.1016/j.appet.2024.107227>
- Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, 70(1), 203–212. <https://doi.org/10.1007/s11336-003-1119-8>
- Zhelyazkova, A., Egger, C., Crushing, R., & Zdrade, S. (2023). *Detecting Democratic Backsliding in Assessment Reports Using Computational Social Science Tools*. ECPR.

Appendix

Chapter 1

Study (Author, Year, Study ID)	Topic of Choice	Mediator	Sample Size	Proportion Mediated	p-value or p-value level	Absolute Indirect Effect	Publication Status
Appelt et al. (2011) Study 1-gains	Intertemporal	Combined	227	0.67	0.004	0.10	1
Lerner et al. (2012) Study 2	Intertemporal	SMRD	189	0.56	0.01	0.14	1
Nichols et al. (2022) Study 1	Intertemporal	SMRD	143	0.29	0.05	0.03	0
Hardisty et al. (2013) Study 1-gains	Intertemporal	Proportion	96	0.18	0.02	0.05	1
Appelt et al. (2011) Study 1-losses	Intertemporal	Combined	380	0.16	0.03	0.05	1
Hardisty et al. (2013) Study 1-losses	Intertemporal	Proportion	96	0.13	0.03	0.05	1
Hardisty et al. (2010) Study 2c	Multi-attribute	Proportion	69	0.46	0.032	0.23	1
Shang et al. (2022) Study 5	Multi-attribute	SMRD	123	0.24	0.05	0.07	1
Spälti et al. (2019) Study 1	Multi-attribute	SMRD	203	0.22	0.000001	6.00	0
Majd et al. (2017)	Multi-attribute	SMRD	102	0.21	0.286	0.05	1
Hardisty et al. (2010) Study 2c	Multi-attribute	SMRD	69	0.18	0.388	0.23	1
Poon et al. (2023) Study 1b	Multi-attribute	SMRD	127	0.10	0.22	0.03	1
Johnson et al. (2007) Study 1	Multi-attribute	SMRD	80	0.02	0.62	0.07	1
Payne et al. (2013) Study 3	Multi-attribute	SMRD	390	0.01	0.51	0.24	1
Wall et al. (2020) Study 3	Risky Choice	Combined	241	0.53	0.05	0.15	1

Table A1.1. Studies included in Meta-Analysis 2.

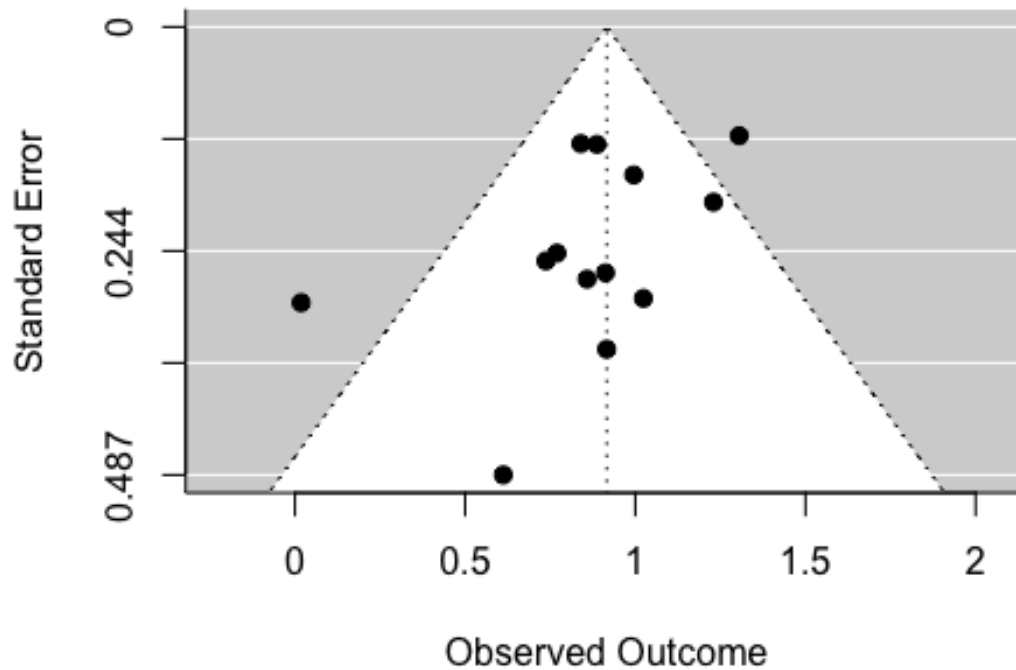


Figure A1.1. Funnel Plot for Meta-Analysis 3a.

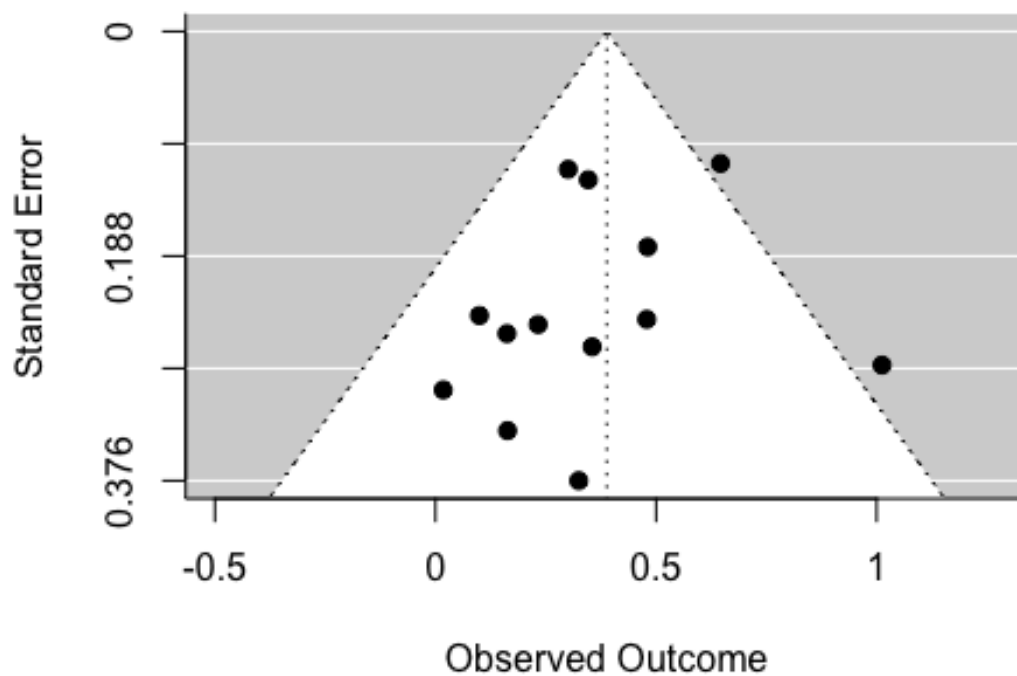


Figure A1.2. Funnel Plot for Meta-Analysis 3b.

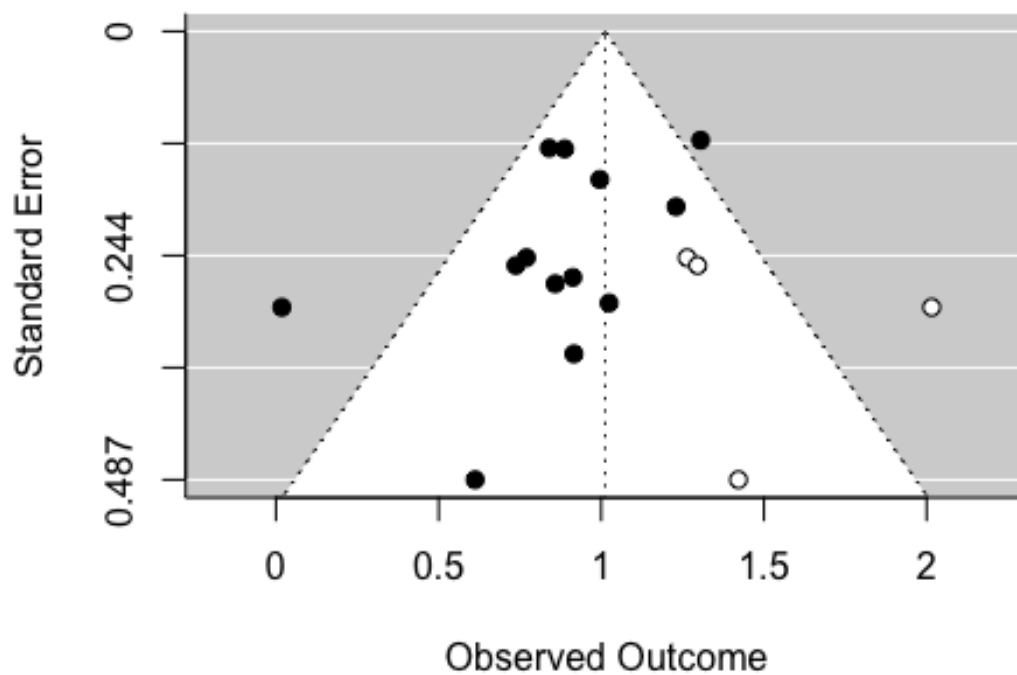


Figure A1.3. Trim & Fill Funnel Plot for Meta-Analysis 3a.

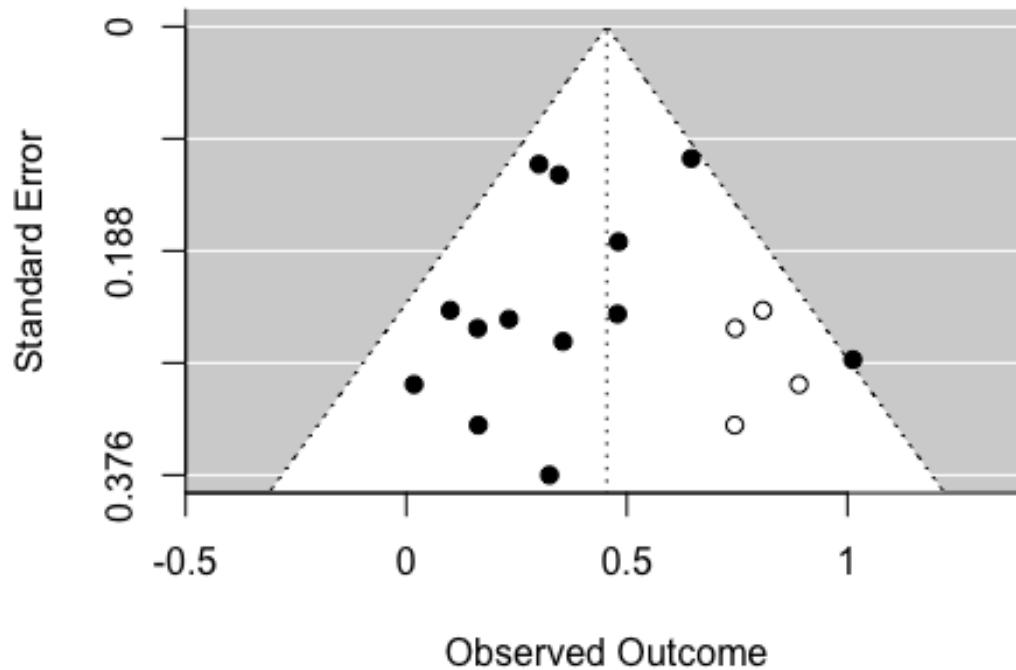


Figure A1.4. Trim & Fill Funnel Plot for Meta-Analysis 3b.

Method specification	Meta-regression results
REML (no moderators)*	$d = 0.34$, $CI95 = [0.27, 0.41]$, $p < .001$, $I^2 = 59.63\%$
REML with choice type as moderators	$d = 0.30$, $CI95 = [0.22, 0.39]$, $p < .001$, $I^2 = 52.61\%$
Fixed effects	$d = 0.34$, $CI95 = [0.27, 0.35]$, $p < .001$, $I^2 = 61.21\%$
REML + robust variance estimation & small sample correction	$d = 0.34$, $CI95 = [0.27, 0.41]$, $t = 9.45$, $p < .001$.

Table A1.2. Alternative model specifications for Meta-Analysis 1. The asterisk (*) indicates the model used in the paper.

Method specification	Meta-regression results
REML (moderator for mediator variable) *	$d = 0.55$, $CI95 = [0.31, 0.98]$, $p = .043$, $I^2 = 25.20\%$
REML with additional moderator for choice type	$d = 0.52$, $CI95 = [0.21, 1.27]$, $p = .15$, $I^2 = 23.54\%$
Fixed effects	$d = 0.64$, $CI95 = [0.48, 0.87]$, $p = .0036$, $I^2 = 0\%$
REML + robust variance estimation & small sample correction	$d = 0.55$, $CI95 = [0.53, 0.58]$, $p = .022$

Table A1.3. Alternative model specifications for Meta-Analysis 2. The asterisk (*) indicates the model used in the paper.

Method specification	Natural order meta-regression results	Unnatural order meta-regression results
----------------------	---------------------------------------	---

REML (no moderators)*	$d = 0.92$, CI95 = [0.74, 1.09], $p < .001$, $I^2 = 49.24\%$	$d = 0.39$, CI95 = [0.25, 0.53], $p < .001$, $I^2 = 27.18\%$
REML with frame type as moderators	$d = 0.85$, CI95 = [0.22, 0.39], $p < .001$, $I^2 = 34.19\%$	$d = 0.39$, CI95 = [0.09, 0.69], $p < .001$, $I^2 = 15.74\%$
Fixed effect	$d = 0.96$, CI95 = [0.85, 1.08], $p < .001$, $I^2 = 49.51\%$	$d = 0.40$, CI95 = [0.29, 0.52], $p < .001$, $I^2 = 25.68\%$
REML + robust variance estimation & small sample correction	$d = 0.92$, CI95 = [0.75, 1.08], $t = 11.03$, $p < .001$	$d = 0.39$, CI95 = [0.25, 0.52], $t = 5.64$, $p < .001$

Table A1.4. Alternative model specifications for Meta-Analysis 3. The asterisk (*) indicates the model used in the paper.

Effect	Pooled Effect Size	95% Confidence Interval	I ² (%)
None (full)	0.55	[0.31, 0.98]	25.2%
Effect size 1	0.41	[0.15, 1.08]	25.4%
Effect size 2	0.62	[0.35, 1.10]	23.5%
Effect size 3	0.53	[0.23, 1.22]	27.8%
Effect size 4	0.56	[0.32, 0.97]	23.8%
Effect size 5	0.55	[0.31, 0.98]	26.9%
Effect size 6	0.55	[0.31, 0.98]	25.9%
Effect size 7	0.55	[0.31, 0.98]	25.6%
Effect size 8	0.55	[0.31, 0.98]	26.8%
Effect size 9	0.63	[0.48, 0.83]	0%
Effect size 10	0.55	[0.31, 0.98]	27.0%
Effect size 11	0.55	[0.31, 0.98]	26.9%
Effect size 12	0.55	[0.30, 1.00]	28.7%
Effect size 13	0.55	[0.30, 0.99]	27.8%
Effect size 14	0.63	[0.48, 0.83]	0%
Effect size 15	0.55	[0.31, 0.98]	26.8%

Table A1.5. Leave-one-out sensitivity analysis for Meta-Analysis 2. Each row represents a meta-regression with one study missing.

Chapter 2

Characteristic	Overall, N = 3,755 [†]	US CCS-Adjacent Sample, N = 1,820 [†]	US Representative Sample, N = 1,935 [†]
Age	M = 47 (SD = 17)	M = 42 (SD = 16)	M = 51 (SD = 17)
HH Income			
None	154 (4.1%)	93 (5.1%)	61 (3.2%)
Less than \$20,000	788 (21%)	402 (22%)	386 (20%)
\$20,000 - \$40,000	987 (26%)	467 (26%)	520 (27%)
\$40,000 - \$80,000	1,017 (27%)	477 (26%)	540 (28%)
\$80,000 - \$120,000	439 (12%)	205 (11%)	234 (12%)
\$120,000 - \$200,000	275 (7.3%)	128 (7.0%)	147 (7.6%)
More than \$200,000	95 (2.5%)	48 (2.6%)	47 (2.4%)
Education			
Some schooling, but no diploma or degree	197 (5.2%)	128 (7.0%)	69 (3.6%)
High school diploma or GED	1,093 (29%)	538 (30%)	555 (29%)
Some college	1,140 (30%)	596 (33%)	544 (28%)
College degree	883 (24%)	375 (21%)	508 (26%)
Some graduate school	93 (2.5%)	44 (2.4%)	49 (2.5%)
Graduate degree	349 (9.3%)	139 (7.6%)	210 (11%)
Gender			
Female	2,151 (57%)	1,174 (65%)	977 (50%)
Male	1,560 (42%)	613 (34%)	947 (49%)
Non-binary / third gender	32 (0.9%)	22 (1.2%)	10 (0.5%)
Prefer not to say	12 (0.3%)	11 (0.6%)	1 (<0.1%)
Political Orientation			
Very conservative	426 (11%)	160 (8.8%)	266 (14%)
Conservative	538 (14%)	253 (14%)	285 (15%)
Somewhat conservative	436 (12%)	240 (13%)	196 (10%)
Moderate	1,516 (40%)	782 (43%)	734 (38%)
Somewhat liberal	347 (9.2%)	173 (9.5%)	174 (9.0%)
Liberal	290 (7.7%)	141 (7.7%)	149 (7.7%)
Very liberal	202 (5.4%)	71 (3.9%)	131 (6.8%)
Political Affiliation			
Democrat	1,131 (30%)	483 (27%)	648 (33%)
Independent	1,064 (28%)	519 (29%)	545 (28%)
Republican	1,160 (31%)	542 (30%)	618 (32%)
Other	400 (11%)	276 (15%)	124 (6.4%)

[†] M = Mean (SD = SD); n (%)

Table A2.1. Demographic description of Study 1 sample.

Demographic	Question	Response Options
Age	How old are you? Answer with the number of years.	<i>Text entry</i>

Household income	During 2023, what was your approximate yearly household income before tax?	None, Less than \$20,000, \$20,000 - \$40,000, \$40,000 - \$80,000, \$80,000 - \$120,000, \$120,000 - \$200,000, More than \$200,000
Level of education	What is the highest level of education you have attained?	Some schooling, but no diploma or degree, High school diploma or GED, Some college, College degree, Some graduate school, Graduate degree
Gender	What is your gender?	Male, Female, Non-binary / third gender, Prefer not to say
Political Orientation	Which best describes your political orientation?	Very conservative, Conservative, Somewhat conservative, Moderate, Somewhat liberal, Liberal, Very liberal
Political Affiliation	What best describes your political affiliation?	Republican, Democrat, Independent, Other
Race	What is your race/ethnicity? (check all that apply)	American Indian, East Asian, Pacific Islander, Black or African American, White or Caucasian, Hispanic or Latinx, South Asian, Middle Eastern, Other [<i>self-describe</i>]

Table A2.2. Demographic questions in Study 1 & Study 2.

<https://docs.google.com/document/d/1SirNLVN0fFSlDS9fMdyIzr13qCcAOIX6ZoLX2dFKpFM/e dit?tab=t.gh6do8te8glz>

Table A2.3. Coding manual for Study 1.

Topic	Norm Type	Neutral control			Agree frame			Disagree frame		
		Mean	SE	n	Mean	SE	n	Mean	SE	n
Economic priorities	Descriptive	4.33	0.040	910	4.48	0.072	249	3.92	0.090	237
Economic priorities	Injunctive	3.04	0.053	865	2.72	0.099	254	2.61	0.092	250
Environmental practices	Descriptive	3.87	0.038	905	3.79	0.081	245	3.02	0.073	249
Environmental practices	Injunctive	4.49	0.036	894	4.36	0.072	239	3.66	0.086	241

Table A2.4. Descriptive statistics for norm perceptions of the four statements (topic x norm type). The mean, standard error, and number of responses are shown for each frame (neutral control, agree frame, and disagree frame). 1= almost everyone disagrees, 3= some people disagree, 4=some people agree, 6 = almost everyone agrees.

Norm Topic	Norm Type	Neutral vs. Agree		Neutral vs. Disagree		Agree vs. Disagree	
		U	p	U	p	U	p
Economic priorities	Descriptive	121312	0.075	124588	0.00013	36030	<0.0001
Economic priorities	Injunctive	96573	0.0028	125147	0.00011	32641	0.58

Environmental practices Descriptive	106880	0.36	159084	<0.0001	41278	<0.0001
Environmental practices Injunctive	99434	0.084	146122	<0.0001	37322	<0.0001

Table A2.5. Summary table showing the results of a series of Mann-Whitney tests. The conditions being compared are noted in the column title, and U- and p-statistics are reported.

Characteristic	N = 540 [†]
Age	
18-24 years old	80 (15%)
25-34 years old	96 (18%)
35-44 years old	94 (17%)
45-54 years old	82 (15%)
55-64 years old	120 (22%)
65+ years old	68 (13%)
HH Income	
Less than \$25,000	75 (14%)
\$25,000-\$49,999	113 (21%)
\$50,000-\$74,999	113 (21%)
\$75,000-\$99,999	91 (17%)
\$100,000-\$149,999	87 (16%)
\$150,000 or more	51 (9.4%)
Prefer not to say	10 (1.9%)
Gender	
Female	272 (50%)
Male	259 (48%)
Non-binary / third gender	6 (1.1%)
Prefer not to say	3 (0.6%)
Political Affiliation	
Democrat	169 (31%)
Independent	197 (36%)
No preference	8 (1.5%)
Other	3 (0.6%)
Republican	163 (30%)
[†] n (%)	

Table A2.6. Demographics for Study 2.



We want you to consider whether or not most people stay awake until midnight on New Year's Eve. Think about your own experiences, the experiences of people you know, and representations in media. Type out the memories that go through your head as you consider this question.

Do most people stay awake until midnight on New Year's Eve?

Please tell us what goes through your mind, one memory at a time, as you consider your answer.

List all memories that come to mind as you consider the different ways you and people you know celebrate New Year's Eve. Carefully consider the question, we want to hear all the different perspectives you have. Please enter each memory separately and one at a time below.

Please type each of your **complete memories as it comes to mind** in the box below. After entering a thought, hit the **"enter" or "return" key on your keyboard to submit** your thought.

Each thought should be in its own box. After you hit enter/return, a new box will appear for your next thought. **Please try to list at least 4 thoughts.**

Figure A2.1. Practice task for Study 2.

Chapter 3

Appendix I: Pilot analysis investigating whether GPT can identify social norm information in text

Prior to testing GPT models on the main text corpus, I ran a pilot on an earlier set of data from the same corpus to (1) examine feasibility of the testing pipeline, (2) test the prompts and parameters used in the main analysis, and (3) leverage a time-frozen model of GPT.

Methods

Expert Coding. Expert raters are considered those with formal and extended training on social norms, both their definition and the behavioral science literature associated. Two research assistants (RAs) coded all 3,903 segments in the evaluation set. The RAs received training on social norms by reading academic papers, participating in a series of discussions with the lead authors, and practicing coding. The RAs tested their knowledge by independently coding X segments and discussing their characterization of each segment with lead authors. During the education process, inter-rater reliability increased from $ICC = 0.66$ to $ICC = 0.80$. The inter-rater reliability is lower than acceptable for text categorization tasks. This is expected for the reasons discussed in the introduction—norms in real-world communication are often implied or contextual. For this reason, each segment was coded by at least two RAs and all disagreements were resolved through discussion and consultation with a lead author.

Each segment in the evaluation set was coded by two RAs and the segments were viewed in a randomized order. The two RAs first rated each batch of ~300 segments independently from each other. Segments were coded for whether “the sentence contains or communicates a social norm”, 1 = no norm present, 2 = a descriptive norm is present, 3 = an injunctive norm is present, 4 = both a descriptive and injunctive norm are present, 5 = there is not enough information to determine whether there is a norm present. After each set of 300 segments was coded, the RAs compared codes and reconciled any disagreement in coding through discussion. Next, each set of segments was passed to one of the lead authors. Before viewing the RA codes, one of the two lead authors coded a random 20% of all segments. Then, an author (JC or KN) reviewed every segment that the RAs disagreed on in their original coding and assigned a final code based on the author’s evaluation of the segment and the RAs reconciled code. The authors also compared all segments in the random 20% that were coded to the RA codes to assess RA accuracy. When there was disagreement, the author reviewed the segment and made a final decision; when there were patterns of disagreement, the author communicated directly with the RAs to update their training.

Novice Rater Comparison Set. 900 segments were additionally coded by MTurk workers. Each worker passed a qualification task similar to the one administered in the Prolific survey. The qualification survey gave participants a brief introduction into social norms and the different types of norms that can be communicated before giving them the same three segments to code as were given to Prolific participants. After qualifying—by answering two out of three norm categorization questions—workers were given access to up to ten sets of ten norms each. Workers were asked to categorize each text segment for whether “the sentence contains or communicates a social norm”, 1 = no norm present, 2 = a descriptive norm is present, 3 = an injunctive norm is present, 4 = both a descriptive and injunctive norm are present, 5 = there is not enough

information to determine whether there is a norm present. The instructions are identical to what was shown to expert raters and GPT-4.

GPT Coding (untuned): The full evaluation set ($n = 3,903$) was coded by the untuned GPT-4 (model ID: gpt-4-0613) model using the OpenAI API. The model I selected had time-frozen input data, which guaranteed that none of the text I were coding was in the model's training data. The code for this query was written in R and run from a Lenovo ThinkStation P620 ThinkStation Threadripper PRO 3945WX. I used a temperature of 0 to ensure the highest predictability, meaning if the model were run multiple times variation in response would be minimal (Rathje et al., 2024). The classification tasks were run in August 2024.

GPT Coding (fine-tuning). The guidelines from OpenAI require a minimum of ten examples to fine tune and recommend 50 to 100 examples because there are improvements in fine-tuning performance. I created a training set with 100 randomly-selected segments from the expert-coded set and a prompt that matches what was used in the untuned model, see above. The model was validated against a different set of 100 randomly-selected segments from the expert-coded set. The OpenAI fine-tuning UI was used to fine-tune the gpt-4o-2024-08-06 model. I use OpenAI's recommended number of epochs, 3, and batch size, 1, for the training data. Two checkpoints are captured to evaluate over-fitting—after 100 steps and after 200 steps. The gpt-4-0613 model, the un-tuned model version that I used, is not available for fine-tuning (it is listed as available on the OpenAI website, but it is not accessible; some forums list it as 'experimental'). The fine-tuning process results in 0.0004 training loss, 0.0004 validation loss, and 0.2019 full validation loss. All metrics decrease from the first checkpoint to the second checkpoint to the final model.

After evaluating the first fine-tuned model, I ran another round of tuning on that fine-tuned model to improve accuracy in identifying injunctive norms and sentences that contain both injunctive and descriptive. The original training set was randomly selected, but did not include text coded as a 3 or 4, due to very low rates of occurrence in the expert-coded data. Therefore, the second step of fine-tuning included a balanced and smaller sample—10 segments per code (1, 2, 3, 4, 5) and the validation set was half the size with the same distribution. All segments are unique, which means no segment is used in multiple training or validation sets.

Results

How accurate was GPT-4 when identifying norm information?

Compared to expert raters, *GPT-4 untuned* was 64% accurate when identifying whether a segment contains norm information. Using a one-sample t-test, I found that GPT-4 performed significantly better than chance ($t(3902) = 17.609, p < .00001$). Using a signal detection theory framework, GPT-4's hit rate was 71% and the false alarm rate was 39%—indicating that GPT-4 accurately identified 71% of the true norms as norms and identified 39% of the not-norms, falsely, as norms (Table A1). To contextualize these findings, I compared *GPT-4 untuned* to naive human raters and found that *GPT-4 untuned* was significantly more accurate overall ($t(899) = 5.06, p < .0001$). The bias of the two classifiers, which describes how conservative or liberal each classifier is, indicated that *GPT-4 untuned* was a slightly more conservative classifier than naive human raters. Examining the hit and false alarm rates, I saw that increased accuracy and higher conservatism was associated with a small decrease in hit rate and moderate decrease

in false alarm rate. This comparison suggested that *GPT-4 untuned* was better than naive human raters at surfacing norms from naturalistic text.

Classifier	Overall Accuracy	Hit Rate	False Alarm Rate	Sensitivity (A)	Bias (beta) High = more conservative
GPT-4 untuned	64% ***	71%	39%	0.72	0.84
GPT4o fine-tuned	71% ***	51%	22%	0.71	1.57
Novice human raters (MTurk)	56% ***	77% (SD = .14)	48% (SD = .22)	0.74 (SD = .05)	0.77 (SD = .45)

Table A3.1. Summary of norm surfacing classifier accuracy. Three classifiers are evaluated for their accuracy compared to expert human raters—GPT-4 untuned, GPT4o fine-tuned, and novice human raters. Overall accuracy is the percentage of accurate classifications, regardless of type, and is compared to chance (50% accuracy) using a one sample t-test, significance levels are shown as: *** $p < .0001$, ** $p < .001$, * $p < .01$. Full t-test results are shown in Table A3. Hit rate is calculated as the percentage of true signals that are accurately detected; false alarm rate is calculated as the percentage of noise that is inaccurately classified as a signal. Sensitivity is calculated as Zhang & Mueller’s non-parametric sensitivity (A) and bias is calculated as Zhang & Mueller’s non-parametric bias (b) (<http://shanetmueller.info:3838/aprime/>). Sensitivity describes the standardized mean difference between the signal and noise distributions; bias describes how conservative or liberal the classifier is with higher values indicating a more conservative criterion.

How accurate was GPT-4 when identifying specific types of norm information?

The overall accuracy of *GPT-4 untuned* is lower, 53%, when evaluating the type of norms classification (descriptive, injunctive, or both). *GPT-4 untuned* is more accurate when identifying injunctive norms than descriptive norms ($t(176) = -3.94$, $p = .0001$; Table A2). The sensitivity of identifying injunctive norms is higher than for injunctive norms, suggesting there is more overlap between the signal and noise distributions in the descriptive norm task than in the injunctive one. The descriptive norm classifier task is also more conservative than the injunctive one.

Classification task: <i>Type of norm</i>	Overall Accuracy	Hit Rate	False Alarm Rate	Sensitivity (A)	Bias (beta)
Descriptive norm	73%	44%	19%	0.70	1.65
Injunctive norm	77%	62%	22%	0.77	1.34

Table A3.2. Summary of norm surfacing GPT-4 untuned accuracy on different norm types. Signal detection theory metrics comparing GPT-4 untuned and expert classifications of descriptive and injunctive norms. Overall accuracy is the percentage of accurate classification of

the norm type. Hit rate is calculated as the percentage of true signals that are accurately detected; false alarm rate is calculated as the percentage of noise that is inaccurately classified as a signal. Sensitivity is calculated as Zhang & Mueller's non-parametric sensitivity (*A*) and bias is calculated as Zhang & Mueller's non-parametric bias (*b*) (<http://shanetmueller.info:3838/aprime/>). Sensitivity describes the standardized mean difference between the signal and noise distributions; bias describes how conservative or liberal the classifier is with higher values indicating a more conservative criterion.

Model	Accuracy Compared to chance
GPT-4	$t(3902) = 17.609, p < .00001$
GPT4o fine-tuned	$t(999) = 14.712, p < .00001$
Novice human raters (MTurk)	$t(899) = 4.722, p < .00001$

Table A3.3. Results from three one sample t-tests comparing model accuracy to chance (50% accuracy, $\mu = .50$) from the SDT analysis.

Appendix II: Supplemental materials for main GPT evaluation

Classification task: <i>Type of norm</i>	Overall Accuracy	Hit Rate	False Alarm Rate	Sensitivity (<i>A</i>)	Bias (<i>b</i>)
<i>GPT-4o Untuned: 54% overall accuracy</i>					
Descriptive norm	74%	38%	20%	0.66	1.52
Injunctive norm	85%	12%	4%	0.68	2.47
Both descriptive & injunctive	82%	39%	16%	0.70	1.74
<i>GPT-4o Fine-tuned: 74% overall accuracy</i>					
Descriptive norm	86%	51%	8%	0.82	2.24
Injunctive norm	86%	34%	6%	0.78	2.59
Both descriptive & injunctive	95%	0%	0%	<i>n/a</i>	<i>n/a</i>

Table A3.4. Summary of norm surfacing GPT-4 untuned accuracy on different norm types.

Signal detection theory metrics comparing untuned GPT-4o and fine-tuned GPT-4o to expert classifications of descriptive and injunctive norms. Overall accuracy is the percentage of accurate classification of the norm type. Hit rate is calculated as the percentage of true signals that are accurately detected; false alarm rate is calculated as the percentage of noise that is inaccurately classified as a signal. Sensitivity is calculated as Zhang & Mueller's non-parametric sensitivity (*A*) and bias is calculated as Zhang & Mueller's non-parametric bias (*b*). Sensitivity describes the standardized mean difference between the signal and noise distributions; bias

describes how conservative or liberal the classifier is with higher values indicating a more conservative criterion.

<i>Predictors</i>	Is Norm (fine-tuned GPT-4o)			Is Norm (untuned GPT-4o)		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.04	0.03 – 0.06	<0.001	0.29	0.22 – 0.39	<0.001
Channel [MSNBC]	5.96	3.69 – 9.61	<0.001	1.33	0.86 – 2.07	0.201
Channel [FOX]	1.39	0.83 – 2.33	0.205	2.59	1.72 – 3.91	<0.001
Week	1.05	1.04 – 1.06	<0.001	1.03	1.02 – 1.04	<0.001
Sentiment	2.91	2.27 – 3.74	<0.001	1.99	1.61 – 2.46	<0.001
Channel [MSNBC] x Week	0.97	0.96 – 0.98	<0.001	1.00	0.99 – 1.02	0.479
Channel [FOX] x Week	0.98	0.97 – 1.00	0.011	0.98	0.97 – 0.99	<0.001
Channel [MSNBC] x Sentiment	2.92	2.09 – 4.08	<0.001	0.91	0.69 – 1.20	0.505
Channel [FOX] x Sentiment	0.44	0.33 – 0.60	<0.001	0.26	0.20 – 0.34	<0.001
Observations	157754			157754		
R ² Tjur	0.028			0.010		

Table A3.5. Two binomial logistic regression predicting norm prominence in text segments for each of the classifiers. The fine-tuned GPT-4o model results are on the left and the untuned GPT-4o model results are on the right. The model regresses the binary norm indicator on channel, week, sentiment, the interaction between channel and week, and the interaction between channel and sentiment. The reference level for the channel factor is CNN.

Appendix III: Extended results for descriptive analysis

This section contains supplementary materials for the descriptive analysis presented in the main chapter.

Topic	Terms
Ignoring Court Decisions	Supreme Court, SCOTUS, superior court, county court, federal court, department of justice, court of appeals, court, courts, courtroom, judge, federal judge, jury, grand jury, prosecutor, prosecution, defendant, defendants, indictment, charges, overturn the ruling, investigate, investigation, attorney, ruling, court decision, rule of law, held accountable, independence of the judiciary, state court, judicial activism, judiciary, clear legal, clear law, checks and balances, independent judiciary
Reducing Polling Stations	vote, vote counting, voter ID, voter registration, voter suppression, voting, voting booth, polling station, polling place, stay in line, precinct, voting registration, register to vote, gerrymandering, early voting

Use of Executive Orders	executive order, executive orders, executive power, presidential action, presidential actions, mandate, Federal Register
Media Censorship	censor, censorship, news coverage, media coverage, news attention, media attention, suppression of speech, public communication, freedom of speech, misinformation, disinformation, speech rights, government censorship, protected speech
Party Loyalty	stop the steal, January 6th, party loyalty, loyal to, insurrectionists, insurrection, presidential transition, transition of power, general services administration, republican loyalty, democratic loyalty

Table A3.6. Topic dictionaries. List of keywords used to identify each of the democracy topics.

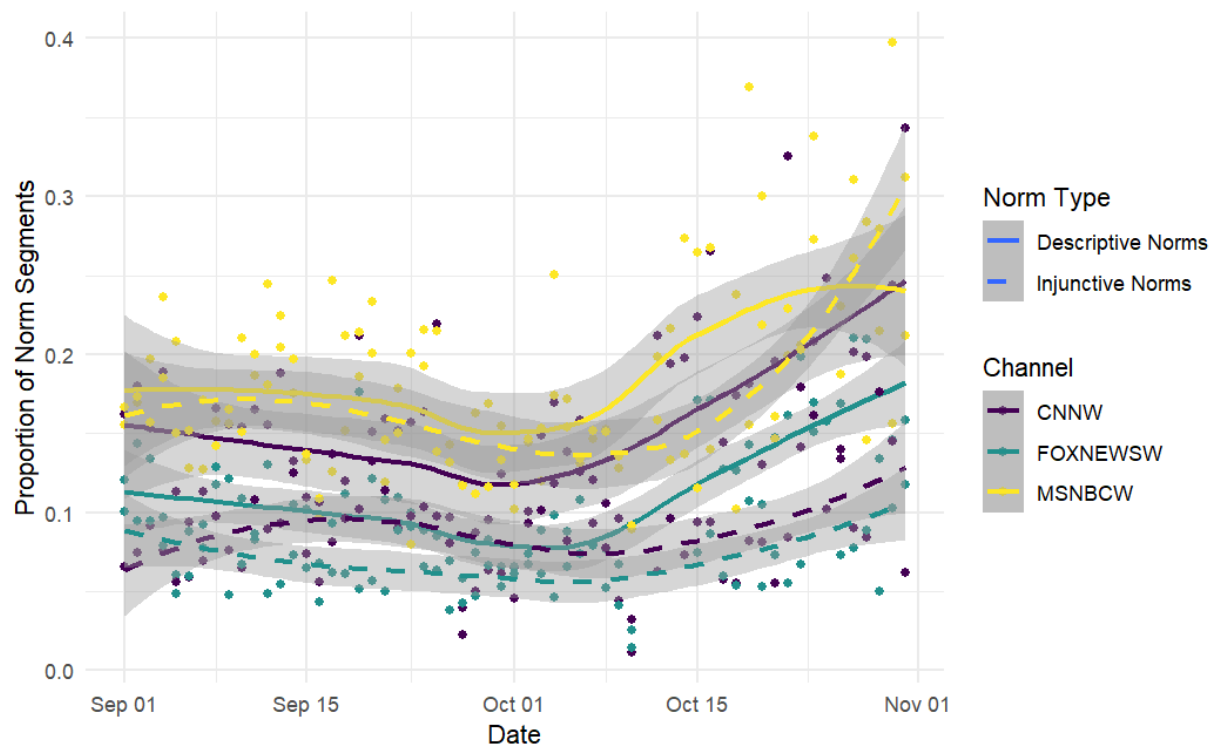


Figure A3.1. The proportion of norm segments is plotted over time. Proportion of descriptive norms (solid line) and injunctive norms (dashed line) is calculated for each day on each channel. Channel is shown in color: CNN in purple, Fox in teal, and MSNBC in yellow.

Descriptive and injunctive norm prevalence in television news

Descriptive norms were more common than injunctive norms in the television news corpus; descriptive norms appeared in 14.2% of text segments and injunctive norms appeared in 10.2% ($X^2(1) = 1152.7, p < 0.001$). This pattern held for all three channels. On CNN 14.7% of segments contained descriptive norms and 8.8% contained injunctive norms ($X^2(1) = 1019.8, p < 0.001$). On Fox 10.4% of segments contained descriptive norms and 6.7% contained injunctive norms ($X^2(1) = 459.6, p < 0.001$). And on MSNBC 17.7% of segments contained descriptive

norms and 16.0% contained injunctive norms ($\chi^2(1) = 45.43, p < 0.001$). Using two binomial logistic regression analysis, I examined how descriptive and injunctive norm prominence changed over time and by sentiment for each channel. Table A6 shows the results of the two models predicting if a segment contains a descriptive norm (left) and if a segment contains an injunctive norm (right). In the descriptive norms model, channel interacted with week and sentiment (Figures A2a & A3a). As text sentiment increased (became more positive) on MSNBC, it was more likely to contain norm information, compared to CNN ($OR = 2.05, p < 0.001$). The interaction between channel and week indicated that while all channels had a higher probability of containing a descriptive norm in later weeks, the rate of probability increase was greatest on MSNBC. Taken together, the way MSNBC communicated norms differed from CNN and Fox—sentiment predicted descriptive norms only on MSNBC. In the injunctive norms model, channel significantly interacted with week and sentiment (Figures A2b & A3b). The significant interaction between sentiment and channel indicated that, like in the ‘is norm’ model, more positive sentiment is associated with a higher likelihood of communicating an injunctive norm on MSNBC and CNN, but the probability increased more per unit for MSNBC than for CNN ($OR = 1.78, p = 0.014$). For Fox, increasing sentiment is associated with less change in norm probability than for CNN ($OR = 0.23, p < 0.001$). The norm probability increased slightly for CNN over time and decreased for FOX ($OR = 0.97, p = 0.001$) and MSNBC ($OR = 0.97, p = 0.001$).

<i>Predictors</i>	Is Descriptive Norm (fine-tuned GPT-4o)			Is Injunctive Norm (fine-tuned GPT-4o)		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.02	0.01 – 0.03	<0.001	0.04	0.02 – 0.06	<0.001
Channel [MSNBC]	3.21	1.80 – 5.75	<0.001	5.92	3.04 – 11.53	<0.001
Channel [FOX]	0.72	0.39 – 1.34	0.304	3.07	1.41 – 6.68	0.005
Week	1.06	1.05 – 1.07	<0.001	1.02	1.01 – 1.03	0.001
Sentiment	1.06	0.79 – 1.43	0.697	8.80	6.14 – 12.61	<0.001
Channel [MSNBC] x Week	0.98	0.96 – 0.99	0.001	0.97	0.96 – 0.99	0.001
Channel [FOX] x Week	1.00	0.98 – 1.01	0.894	0.97	0.95 – 0.99	0.001
Channel [MSNBC] x Sentiment	2.05	1.39 – 3.03	<0.001	1.78	1.12 – 2.82	0.014
Channel [FOX] x Sentiment	0.88	0.61 – 1.27	0.479	0.23	0.14 – 0.35	<0.001
Observations	157754			157754		
R ² Tjur	0.009			0.020		

Table A3.7. Two binomial logistic regression predicting descriptive norm prominence (left) and injunctive norm prominence (right) in text segments. The models regress the respective binary norm indicator on channel, week, sentiment, the interaction between channel and week, and the interaction between channel and sentiment. The reference level for the channel factor is CNN.

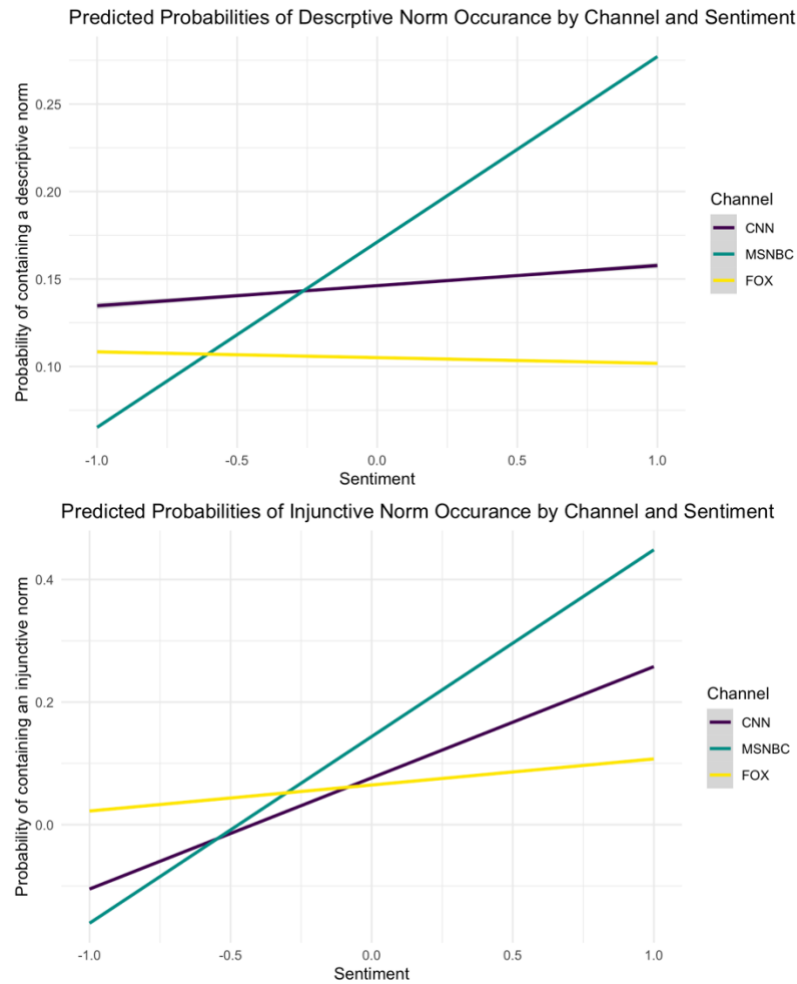


Figure A3.2a & A3.2b. Two interaction plots for the binomial logistic regression models in Table A7—descriptive norm on the left and injunctive norm on the right. The plot shows how the relationship between sentiment (x-axis) and probability of text containing a norm (y-axis) varies by channel (color; purple = CNN, teal = MSNBC, yellow = Fox).

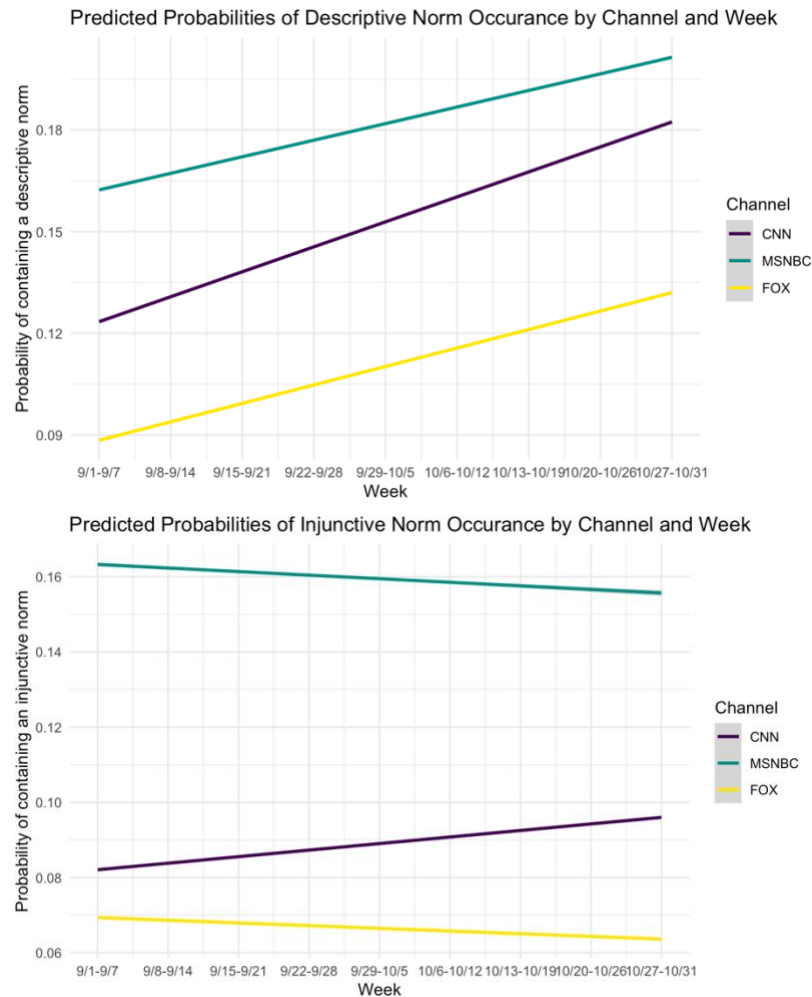


Figure A3.3a & A3.3b. Two interaction plots for the binomial logistic regression models in Table A7—descriptive norm on the left and injunctive norm on the right. The plot shows how the relationship between week (x-axis) and probability of text containing a norm (y-axis) varies by channel (color; purple = CNN, teal = MSNBC, yellow = Fox).

Topic Coverage of Democratic Norm Violations			
<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.32	0.29 – 0.35	<0.001
Sentiment	0.31	0.24 – 0.40	<0.001
MSNBC	1.29	1.25 – 1.34	<0.001
Fox News	0.88	0.85 – 0.91	<0.001
Sentiment x MSNBC	5.03	3.62 – 6.99	<0.001
Sentiment x Fox News	1.38	1.02 – 1.86	0.039

Random Effects

σ^2	3.29
$\tau_{00 \text{ date}}$	0.14
ICC	0.04
N _{date}	60
Observations	157754
Marginal R ² / Conditional R ²	0.012 / 0.052

Table A3.8. Mixed effects logistic regression model where topic coverage of democratic norms was regressed on sentiment and channel, with an interaction and random effect for date. The reference group for channel was CNN.

<i>Predictors</i>	Ignoring Court Decisions(binary)			Reducing Polling Stations(binary)			Use of Executive Orders(binary)			Media Censorship(binary)			Party Loyalty(binary)		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.13	0.11 – 0.15	<0.001	0.12	0.10 – 0.14	<0.001	0.00	0.00 – 0.00	<0.001	0.01	0.01 – 0.02	<0.001	0.00	0.00 – 0.00	<0.001
SentimentGI	0.23	0.16 – 0.32	<0.001	1.21	0.86 – 1.70	0.268	0.10	0.02 – 0.64	0.014	0.07	0.03 – 0.16	<0.001	0.25	0.05 – 1.19	0.082
Channel [MSNBC]	1.21	1.16 – 1.26	<0.001	1.24	1.18 – 1.30	<0.001	0.73	0.55 – 0.97	0.029	1.28	1.15 – 1.42	<0.001	5.83	4.94 – 6.89	<0.001
Channel [FOX]	0.94	0.90 – 0.98	0.005	0.83	0.79 – 0.87	<0.001	1.95	1.58 – 2.42	<0.001	0.96	0.86 – 1.08	0.502	1.03	0.84 – 1.26	0.806
SentimentGI x Channel [MSNBC]	3.45	2.28 – 5.21	<0.001	4.59	2.92 – 7.21	<0.001	19.41	1.38 – 272.66	0.028	5.53	2.11 – 14.53	0.001	2.11	0.41 – 10.98	0.373
SentimentGI x Channel [FOX]	1.38	0.94 – 2.02	0.097	0.82	0.54 – 1.23	0.331	3.74	0.52 – 26.84	0.190	7.40	2.79 – 19.60	<0.001	1.41	0.24 – 8.28	0.701
Random Effects															
σ^2	3.29			3.29			3.29			3.29			3.29		
τ_{00}	0.26 _{date}			0.35 _{date}			0.24 _{date}			0.43 _{date}			0.61 _{date}		
ICC	0.07			0.10			0.07			0.12			0.16		
N	60 _{date}			60 _{date}			60 _{date}			60 _{date}			60 _{date}		
Observations	157754			157754			157754			157754			157754		
Marginal R ² / Conditional R ²	0.008 / 0.082			0.013 / 0.108			0.045 / 0.110			0.012 / 0.127			0.149 / 0.281		

Table A3.9. Results from five mixed effects logistic regression models where topic coverage of each democratic norm was regressed on sentiment and channel, with an interaction and random effect for date. The reference group for channel was CNN.

Specific topic coverage also varied across channels. Using logistic mixed effects regression, I regressed topic prominence on channel and sentiment, interacted, with a random intercept for date. The full model results are shown in Table A9. Four of the five models had significant interactions between sentiment and channel (Figure A4). Text was more likely to be about ignoring court decisions on CNN and Fox when there was negative sentiment. On MSNBC, text was more likely to be about reducing polling stations when there was higher sentiment. Use of executive orders was more likely to occur on CNN and Fox when sentiment was negative. On all channels, media censorship was more likely to be discussed when there was negative sentiment. Finally, the only significant predictor of party loyalty topic prominence was MSNBC. Generally, the democratic norm topics were more likely to be mentioned with negative sentiment, with the exception of reducing polling stations and channel-specific differences (Figure A4).

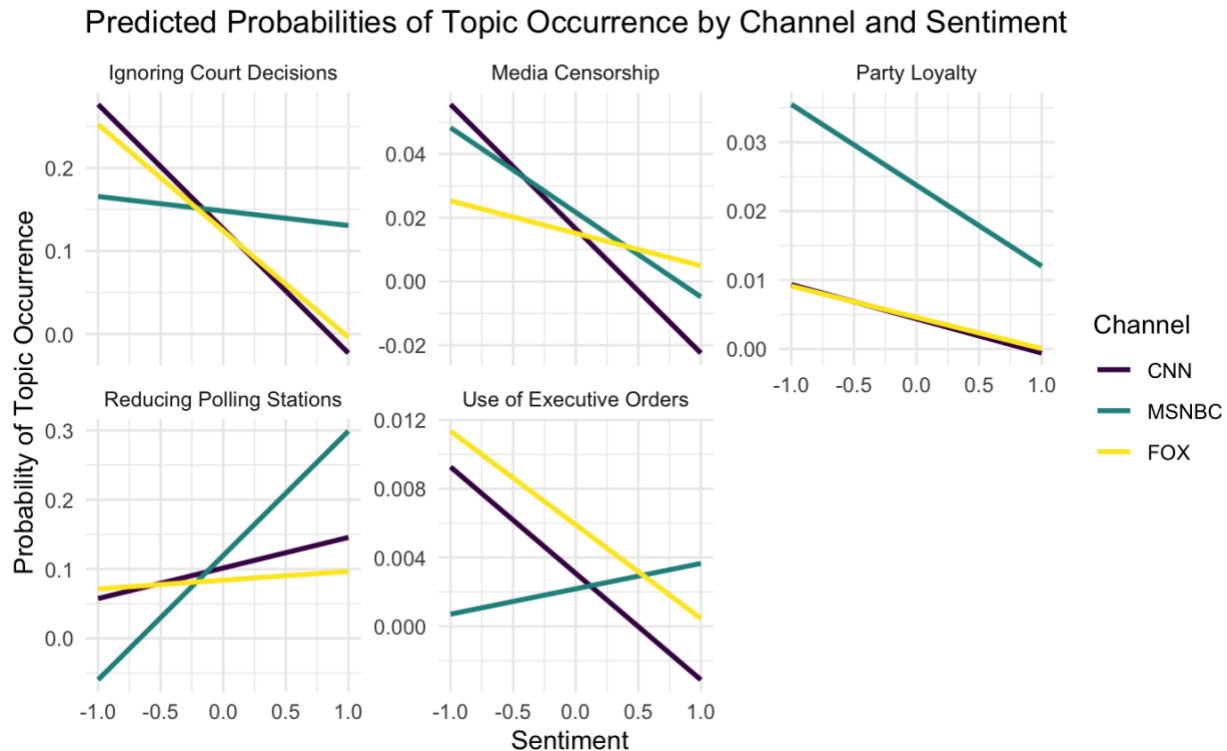


Figure A3.4. Interaction plots for the five binomial logistic regression models predicting democratic topic prominence in text. The full model results are shown in Table A9. For each topic, the plots show the relationship between sentiment (x-axis) and probability of the democracy topic being discussed (y-axis) varies by channel (color; purple = CNN, teal = MSNBC, yellow = Fox). The model includes a random effect for date.

Norm Prominence in Coverage of Democratic Norm Violations			
<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.44	0.39 – 0.50	<0.001
Sentiment	35.40	20.85 – 60.09	<0.001
MSNBC	1.58	1.48 – 1.69	<0.001
Fox News	0.77	0.71 – 0.83	<0.001
Sentiment x MSNBC	1.81	0.89 – 3.69	0.103
Sentiment x Fox News	0.44	0.20 – 0.95	0.036

Random Effects

σ^2	3.29
$\tau_{00 \text{ date}}$	0.18
ICC	0.05
N_{date}	60
Observations	37575

Marginal R^2 / Conditional R^2 0.051 / 0.100

Table A3.10. Results from a mixed effects logistic regression where norm prominence was regressed on sentiment and channel with an interaction and a random effect for date. The reference group for channel was CNN.

<i>Predictors</i>	Norm language in Ignoring Court Decisions Subset			Norm language in Reducing Polling Stations Subset			Norm language in Use of Executive Orders Subset			Norm language in Media Censorship Subset			Norm language in Party Loyalty Subset		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.21	0.18 – 0.24	<0.001	0.88	0.80 – 0.96	0.006	0.40	0.26 – 0.61	<0.001	0.42	0.34 – 0.53	<0.001	0.73	0.51 – 1.04	0.082
SentimentGI	49.53	21.53 – 113.94	<0.001	22.82	10.56 – 49.31	<0.001	79.32	0.89 – 7100.79	0.056	0.52	0.05 – 5.65	0.594	905.45	16.06 – 51034.15	0.001
Channel [MSNBC]	1.45	1.30 – 1.62	<0.001	1.85	1.67 – 2.05	<0.001	1.52	0.79 – 2.94	0.215	1.89	1.49 – 2.41	<0.001	1.17	0.81 – 1.71	0.402
Channel [FOX]	0.84	0.74 – 0.94	0.003	0.78	0.70 – 0.86	<0.001	1.07	0.64 – 1.78	0.801	0.84	0.63 – 1.12	0.237	0.54	0.33 – 0.88	0.014
SentimentGI x Channel [MSNBC]	4.43	1.47 – 13.36	0.008	0.50	0.17 – 1.47	0.205	0.00	0.00 – 2.52	0.088	32.03	1.64 – 623.88	0.022	0.01	0.00 – 0.64	0.031
SentimentGI x Channel [FOX]	0.34	0.10 – 1.12	0.076	0.22	0.07 – 0.70	0.010	0.22	0.00 – 65.66	0.604	12.99	0.45 – 373.00	0.134	0.00	0.00 – 0.32	0.020
Random Effects															
σ^2	3.29			3.29			3.29			3.29			3.29		
τ_{00}	0.17 <i>date</i>			0.05 <i>date</i>			0.01 <i>date</i>			0.17 <i>date</i>			0.13 <i>date</i>		
ICC	0.05			0.02			0.00			0.05			0.04		
N	60 <i>date</i>			60 <i>date</i>			59 <i>date</i>			60 <i>date</i>			59 <i>date</i>		
Observations	19899			16627			559			2543			1548		
Marginal R^2 / Conditional R^2	0.053 / 0.099			0.049 / 0.064			0.017 / 0.020			0.048 / 0.096			0.043 / 0.078		

Table A3.11. Results from five mixed effects logistic regression where norm prominence within coverage for each of the five democratic norms was regressed on sentiment and channel with an interaction and a random effect for date. The reference group for channel was CNN.

To further examine how norm information was used when discussing democratic norm topics, I used logistic mixed effects regressions. Norm presence was regressed on sentiment and channel, interacted, and a random intercept for date. Full regression results in Table A11. There was a significant interaction between channel and sentiment in all models except the executive orders

model. Norm prominence in the news coverages on the use of executive orders did not vary by sentiment or channel. The news coverage on ignoring court decisions was more likely to communicate a norm when sentiment was higher—the interaction between sentiment and channel indicated that the association between positive sentiment and norms was strongest on MSNBC (Figure A5). News coverage on reducing polling stations also had a positive association between sentiment and likelihood of communicating a norm; for this topic, the strongest association was on CNN and the channel most likely to communicate a norm on this topic, regardless of sentiment, was MSNBC (Figure A5). In the coverage of media censorship and party loyalty, the sentiment of the coverage did not have a consistent relationship with norm presence. On MSNBC, positive coverage of media censorship was more likely to communicate a norm than on the other channels (Figure A5). On CNN, positive coverage of party loyalty was more likely to communicate a norm than on the other two channels (Figure A5).

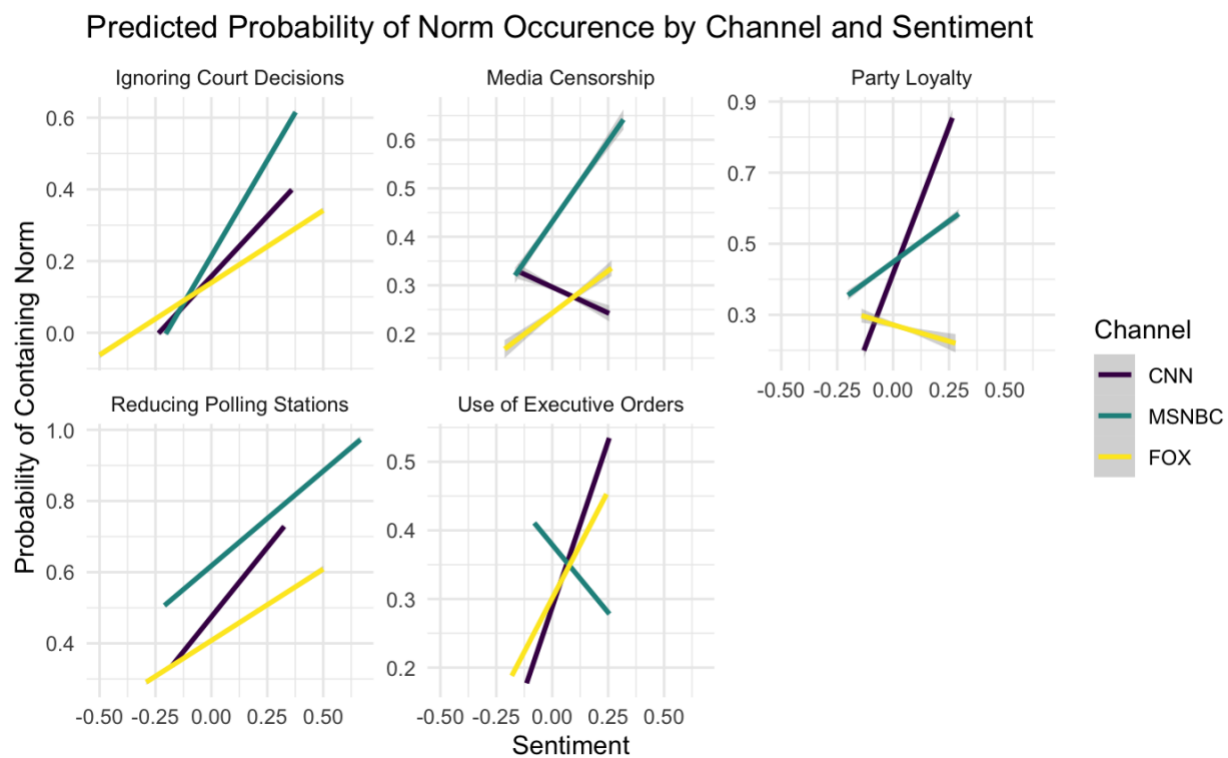


Figure A3.5. Interaction plots for the five mixed effects logistic regressions in Table A11. The interaction between sentiment and channel is shown for each model.

Appendix III: Supplementary material for media diet study

Predictors	Attitude			American Norms			Democrat Norms			Republican Norms		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	0.00	-0.10 – 0.10	0.949	0.00	-0.10 – 0.10	0.974	0.01	-0.09 – 0.11	0.828	-0.00	-0.10 – 0.10	0.951
Topic Coverage	-0.11	-0.27 – 0.04	0.152	-0.02	-0.17 – 0.14	0.846	-0.01	-0.17 – 0.14	0.854	-0.05	-0.20 – 0.10	0.529
Norm Prominence	-0.24	-0.47 – -0.01	0.044	-0.07	-0.30 – 0.17	0.579	-0.22	-0.46 – 0.01	0.061	0.38	0.15 – 0.61	0.001
Sentiment	0.17	-0.06 – 0.40	0.153	-0.03	-0.26 – 0.20	0.804	0.08	-0.16 – 0.31	0.522	-0.27	-0.50 – -0.04	0.020
Observations	378			378			378			378		
R ² / R ² adjusted	0.038 / 0.031			0.011 / 0.003			0.029 / 0.021			0.028 / 0.020		

Table A3.12. Results from four linear regressions predicting attitudes and norm perceptions about democracy from features of media diet. The predictor variables are topic coverage (the proportion of a show that covers democracy topics), norm prominence (the proportion of democracy topic coverage that communicates social norm information), and sentiment (how positive or negative the norm information is).

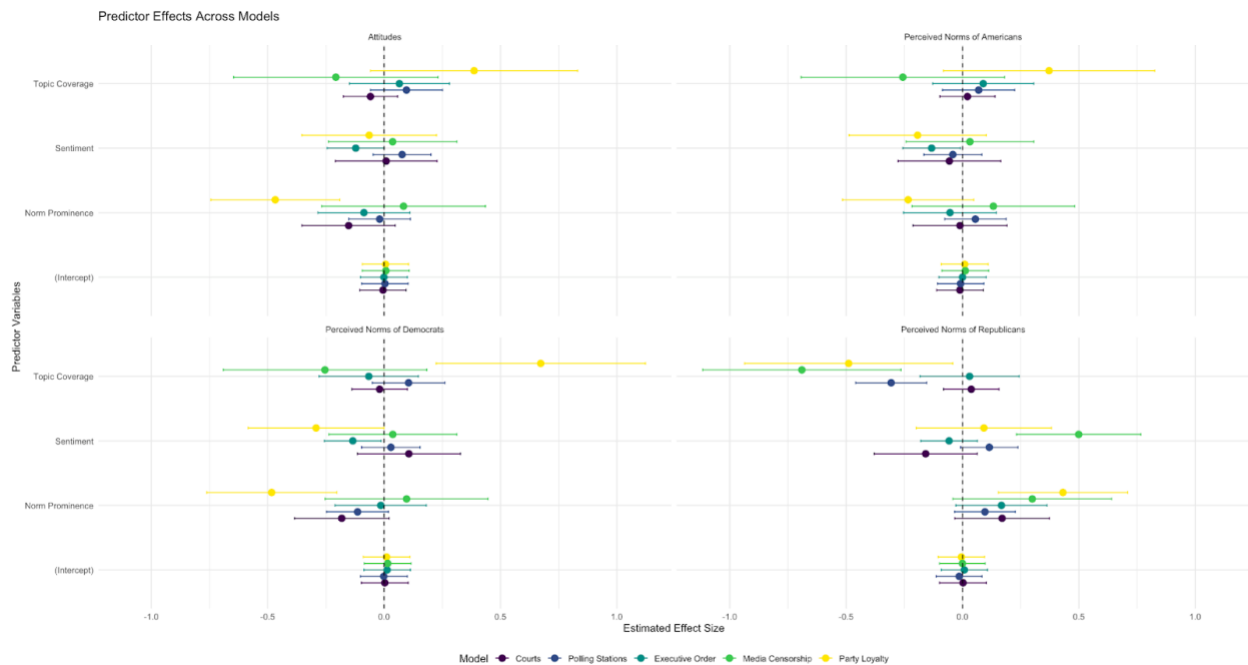


Figure A3.6. Standardized regression coefficients and 95% confidence intervals from linear regression models predicting attitudes (top left), norm perceptions of Americans (top right), norm perceptions of Democrats (bottom left), and norm perceptions of Republicans (bottom right) toward each democratic norm violation (color). Attitudes and norm perceptions were regressed separately on political party, amount of Fox News watched regularly, amount of CNN watched regularly, amount of MSNBC watched regularly, and the interaction between political party and each media diet metric. Political parties are coded as binary variables. The specific norm text is shown in Figure 4.

Predictors	Attitudes			Norm Perceptions of Americans			Norm Perceptions of Democrats			Norm Perceptions of Republicans		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	-0.05	-0.25 – 0.14	0.588	0.11	-0.10 – 0.31	0.312	0.02	-0.18 – 0.21	0.874	0.03	-0.16 – 0.22	0.762
Democrat	0.21	-0.10 – 0.52	0.178	0.09	-0.23 – 0.42	0.576	0.04	-0.27 – 0.34	0.804	0.14	-0.16 – 0.44	0.371
Fox News	0.38	0.17 – 0.59	<0.001	0.19	-0.03 – 0.40	0.094	0.30	0.09 – 0.50	0.005	-0.34	-0.54 – -0.14	0.001
Republican	0.40	0.10 – 0.71	0.009	0.03	-0.29 – 0.35	0.851	0.27	-0.03 – 0.57	0.076	-0.26	-0.56 – 0.04	0.084
CNN	-0.00	-0.22 – 0.22	0.974	0.05	-0.18 – 0.29	0.645	-0.18	-0.40 – 0.04	0.104	0.10	-0.12 – 0.31	0.368
MSNBC	0.05	-0.14 – 0.25	0.599	-0.03	-0.24 – 0.18	0.769	-0.05	-0.24 – 0.15	0.649	0.06	-0.13 – 0.25	0.553
Democrat x Fox News	0.12	-0.17 – 0.41	0.427	0.14	-0.17 – 0.44	0.383	0.16	-0.13 – 0.45	0.275	0.10	-0.19 – 0.38	0.500
Republican x Fox News	-0.51	-0.82 – -0.20	0.001	-0.32	-0.64 – 0.01	0.054	-0.33	-0.64 – -0.03	0.031	0.15	-0.16 – 0.45	0.343
Democrat x CNN	-0.01	-0.31 – 0.28	0.928	-0.06	-0.37 – 0.25	0.698	0.16	-0.13 – 0.45	0.280	-0.01	-0.30 – 0.28	0.921
Republican x CNN	0.01	-0.28 – 0.29	0.958	0.02	-0.28 – 0.32	0.913	0.16	-0.12 – 0.44	0.253	-0.04	-0.32 – 0.24	0.791
Democrat x MSNBC	-0.08	-0.34 – 0.18	0.542	-0.04	-0.31 – 0.23	0.779	-0.09	-0.35 – 0.17	0.490	-0.02	-0.27 – 0.24	0.897
Republican x MSNBC	-0.06	-0.34 – 0.23	0.703	0.07	-0.23 – 0.37	0.654	-0.10	-0.38 – 0.18	0.478	-0.04	-0.31 – 0.24	0.794
Observations	388			388			388			388		
R ² / R ² adjusted	0.130 / 0.104			0.040 / 0.012			0.156 / 0.131			0.170 / 0.146		

Table A3.13. Linear regression models predicting attitudes and norm perceptions. Each DV is regressed on political party, watching Fox, watching CNN, watching MSNBC, and the interaction between political party and each media diet metric. Political party variables are coded as binary.

Table 6: Regression analysis of media diet's relationship to attitudes and norm perceptions of democratic norm violations

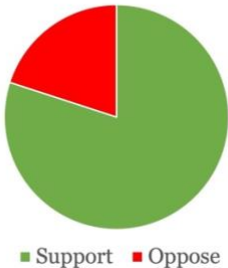
	Dependent variable:									
	Ignoring court decisions		Reducing polling stations		Use of executive orders		Media censorship		Party loyalty	
	Attitude (1)	Norm Perception (2)	Attitude (3)	Norm Perception (4)	Attitude (5)	Norm Perception (6)	Attitude (7)	Norm Perception (8)	Attitude (9)	Norm Perception (10)
Independent	-0.378 (0.649)	21.452 (15.464)	0.007 (0.565)	15.013 (14.853)	-0.247 (0.772)	-2.620 (14.997)	-0.332 (0.741)	5.226 (17.424)	0.529 (0.642)	29.286 (15.765)
Republican	0.526 (0.704)	2.431 (16.770)	0.315 (0.613)	13.563 (16.108)	1.085 (0.838)	3.460 (16.264)	1.157 (0.803)	3.865 (18.895)	1.989* (0.696)	47.542* (17.097)
Fox News	0.405* (0.191)	6.232 (4.544)	0.686* (0.166)	12.824* (4.365)	0.224 (0.227)	4.020 (4.407)	0.605* (0.218)	10.411* (5.120)	0.984* (0.189)	18.720* (4.633)
CNN	-0.155 (0.442)	11.238 (10.541)	-0.227 (0.385)	11.752 (10.124)	-0.140 (0.527)	-8.764 (10.222)	-0.366 (0.505)	-0.561 (11.876)	0.520 (0.438)	9.639 (10.746)
MSNBC	0.111 (0.260)	-1.047 (6.195)	0.237 (0.226)	5.795 (5.951)	0.335 (0.309)	10.553 (6.008)	0.070 (0.297)	-1.026 (6.981)	0.216 (0.257)	6.385 (6.316)
Independent x Fox News	-0.246 (0.296)	-9.413 (7.059)	-0.348 (0.258)	-10.282 (6.780)	-0.253 (0.353)	-1.557 (6.846)	-0.070 (0.338)	-6.231 (7.953)	-0.286 (0.293)	-16.162* (7.196)
Republican x Fox News	-0.741 (0.520)	-4.664 (12.386)	-0.403 (0.453)	-10.743 (11.897)	-0.726 (0.619)	-9.212 (12.012)	-1.453* (0.593)	-8.797 (13.956)	-1.042* (0.514)	-32.089* (12.628)
Independent x CNN	0.234 (0.582)	-18.476 (13.872)	0.417 (0.507)	-8.222 (13.324)	0.353 (0.693)	9.719 (13.453)	0.331 (0.664)	-9.719 (15.630)	-0.491 (0.576)	-11.755 (14.142)
Republican x CNN	0.214 (0.494)	-10.691 (11.783)	0.328 (0.431)	-4.089 (11.318)	-0.062 (0.589)	8.609 (11.428)	0.149 (0.564)	-8.999 (13.277)	-0.779 (0.489)	-11.219 (12.013)
Independent x MSNBC	0.126 (0.359)	2.505 (8.557)	-0.246 (0.313)	-4.052 (8.219)	-0.185 (0.427)	-5.208 (8.299)	-0.121 (0.410)	9.344 (9.642)	-0.010 (0.355)	-9.435 (8.724)
Republican x CNN	-0.136 (0.343)	7.491 (8.165)	-0.292 (0.298)	-8.883 (7.843)	-0.479 (0.408)	-6.772 (7.919)	0.165 (0.391)	10.412 (9.200)	-0.146 (0.339)	-10.588 (8.324)
Constant	2.387* (0.529)	37.611* (12.605)	1.629* (0.461)	11.199 (12.107)	2.081* (0.630)	43.163* (12.224)	2.049* (0.604)	37.709* (14.202)	0.936 (0.523)	20.084 (12.850)
Gender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	382	382	382	382	382	382	382	382	382	382
R ²	0.100	0.101	0.171	0.099	0.089	0.065	0.171	0.115	0.258	0.131
Adjusted R ²	0.026	0.027	0.103	0.025	0.013	-0.012	0.103	0.042	0.197	0.059
Residual Std. Error (df = 352)	1.026	24.456	0.894	23.490	1.222	23.718	1.171	27.555	1.015	24.933
F Statistic (df = 29; 352)	1.347	1.371	2.511*	1.334	1.179	0.841	2.502*	1.575*	4.217*	1.830*

Note:

*p<0.05; **p<0.01; ***p<0.001

Table A3.14. Linear regression models predicting attitudes and norm perceptions for each of the five democratic norm topics. Each attitude and norm perception is regressed on political party, watching Fox, watching CNN, watching MSNBC, the interaction of political party and each media diet metric, and demographics. The reference group for political party is Democrat. Attitudes are measured on a 5-point Likert scale and norm perceptions are measured as the percentage of Americans who support the statement. The specific norm text is shown in Figure 7. Demographics include gender, age, income, education, and geography.

Chapter 4

Stimuli Section	Detail
Policy Introduction	The following is a policy proposal that could happen in most municipalities (i.e., a city or town with a government) across the country. Imagine the scenario below is happening in the place where you currently live. The municipal planning committee where you live is proposing to add a new crosswalk across a highly-trafficked road. There is one lane of traffic in each direction and a high volume of cars on the road. Some pedestrians cross the road where there is no crosswalk rather than going to the next intersection. The road is particularly busy during rush hour when people are driving or walking to work/school. This is a road that you drive on regularly to get to work, and the new crosswalk could increase your commute by 5 minutes. Installing this crosswalk requires a modest amount of public financial support—there is a one-time installation cost and a smaller annual maintenance cost. The planning committee has put this issue on the ballot in an upcoming municipal election. They will ask: Do you support or oppose putting a new crosswalk to cross this road?
Positive Norm Treatment	An article in the municipality's newspaper reports that 80% of people in your neighborhood support adding a crosswalk.  ■ Support ■ Oppose
Negative Norm Treatment	An article in the municipality's newspaper reports that 80% of people in your neighborhood oppose adding a crosswalk.  ■ Support ■ Oppose
Thought lister prompt	Please list the thoughts that go through your mind as you evaluate if you support or oppose putting in a new crosswalk. Carefully consider the pros and cons of this policy.

	Support the new crosswalk: 	Oppose the new crosswalk: 
--	--	--

Table A4.1. Stimuli text and images used for Study 1.

Demographic	Question Text	Response Options
Gender	How do you describe yourself?	Male; Female; Non-binary/third gender; Prefer to self-describe (<i>enter text</i>); Prefer not to say
Age	How old are you?	Under 18; 18-24 years old; 25-34 years old; 35-44 years old; 45-54 years old; 55-64 years old; 65+ years old
Race	Choose one of more races that you consider yourself to be:	(<i>multi-select</i>) White or Caucasian; Black or African American; American Indian/Native American or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; Other; Prefer not to say
Education	What is the highest level of education you have completed?	Some high school or less; High school diploma or GED; Some college, but no degree; Associates or technical degree; bachelor's degree; Graduate or professional degree; Prefer not to say
Household income	What was your total household income before taxes during the past 12 months?	Less than \$25,000; \$25,000 - \$49,999; \$50,000 - \$74,999; \$75,000 - \$99,999; \$100,000 - \$149,999; \$150,000 or more; Prefer not to say
Political Party	Generally speaking, do you usually think of yourself as a Republic, Democrat, Independent, or something else?	Republic; Democrat; Independent; Other (please describe); No preference

Table A4.2. Text of demographic questions.

Causal Mediation Analysis					
Nonparametric Bootstrap Confidence Intervals with the Percentile Method					
	Estimate	95% CI Lower	95% CI Upper	p-value	
ACME	0.289	0.110	0.48	0.0012	**
ADE	0.489	0.232	0.74	0.0002	***
Total Effect	0.779	0.476	1.08	<2e-16	***
Prop. Mediated	0.372	0.169	0.61	0.0012	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Sample Size Used: 295					
Simulations: 10000					

Table A4.3. Mediation results from Study 1: SMRD is investigated as a mediator of the effect of treatment group on policy support.

Causal Mediation Analysis					
Nonparametric Bootstrap Confidence Intervals with the Percentile Method					
	Estimate	95% CI Lower	95% CI Upper	p-value	
ACME	0.418	0.189	0.65	0.0002	***
ADE	0.361	0.143	0.58	0.0012	**
Total Effect	0.779	0.471	1.07	<2e-16	***
Prop. Mediated	0.536	0.321	0.77	0.0002	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Sample Size Used: 295					
Simulations: 10000					

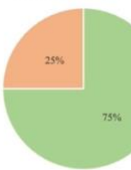
Table A4.4. Mediation results from Study 1: proportion of support thoughts is investigated as a mediator of the effect of treatment group on policy support.

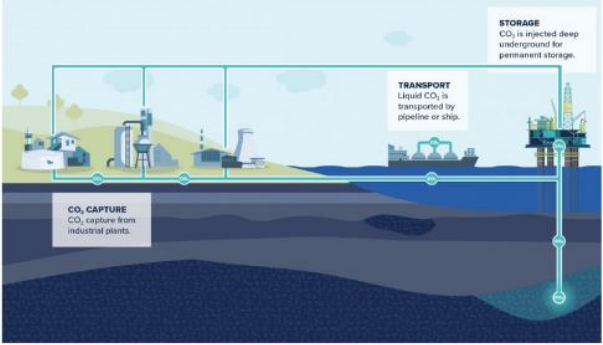
Causal Mediation Analysis					
Nonparametric Bootstrap Confidence Intervals with the Percentile Method					
	Estimate	95% CI Lower	95% CI Upper	p-value	
ACME	0.388	0.177	0.62	<2e-16	***
ADE	0.390	0.158	0.62	0.0012	**
Total Effect	0.779	0.476	1.08	<2e-16	***
Prop. Mediated	0.499	0.283	0.74	<2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Sample Size Used: 295					
Simulations: 10000					

Table A4.5. Mediation results from Study 1: prominence of support thoughts is investigated as a mediator of the effect of treatment group on policy support.

Stimuli Section	Detail
Energy efficiency policy context	We would like to know what you think about a policy that raises energy efficiency standards. We will ask you how much you support the policy and how much money you would be willing to contribute to this policy. Energy-Efficiency Standards The Department of Energy (DOE) establishes energy-efficiency standards on many appliances and products, such as dishwashers, heating and cooling systems, and motor vehicles. Many products that the average person uses every day waste energy due to poor design or cost-cutting measures. Energy-efficiency standards require all household appliances and products to function as usual while keeping energy and water consumption below a specified limit (i.e., a dishwasher that uses fewer gallons of water and kWh of electricity per load). There is a policy proposal to implement mandatory energy efficiency standards to all consumer products - from refrigerators to transportation, including cars. These regulations may cause the price of consumer products to rise, but they would also reduce carbon emissions significantly, thus helping combat climate change. 
Energy efficiency norm treatment	Energy-Efficiency Standards A strong majority of registered voters in the United States think that global warming should be a high priority for the president and Congress. Most Americans say that government regulations are necessary to encourage businesses and consumers to rely more on renewable energy sources. A strong majority of Americans support setting stronger energy efficiency standards for new buildings, appliances, and cars/trucks/SUVs. Most Americans Support Setting Stronger Energy-Efficiency Standards For... New buildings  Appliances  Cars/trucks/SUVs  ■ Support ■ Oppose There is a policy proposal to set higher mandatory energy-efficiency standards for consumer products.
Energy efficiency thought lister prompt	Please carefully consider whether or not you support this policy. Before you indicate your level of support or non-support, please tell us what goes through your mind, one thought at a time, when evaluating this policy. List any thoughts that come to mind about this topic. Please enter your thoughts one at a time below.

Labor organizing policy context	We would like to know what you think about a national policy to give all employees the right to collective bargaining and wage negotiations. We will ask you how much you support the policy and how much money you would be willing to contribute to this policy. Labor Unions & Strikes   Labor unions, or organized groups of workers who unite to make decisions about their work conditions/wages, typically form in response to poor conditions or treatment at work, unfair pay, and workplace inequalities. One way labor unions can communicate discontent and leverage their collective power is through organized strikes, where workers stop labor to communicate their demands and get the attention of the employer. For instance, the recent 2023 Writers Guild of America strike was caused by poor compensation for writers and the negative impact of streaming services on pay and job security. Their goals included increasing pay, job security, and the size of writers' rooms. Across the U.S., there have been more than 200 labor strikes involving more than 320,000 workers in this year, 2023. There is a policy proposal to implement a national law to protect collective bargaining, the process in which workers negotiate their pay and other working conditions/rights. This policy would also help protect the formation of labor unions and strikes.
Labor organizing norm treatment	Labor Unions & Strikes A strong majority of Americans support labor unions and most support workers going on strike to negotiate for better wages, benefits, and working conditions. A strong majority of Americans think it is bad for the U.S. that labor unions have gotten smaller. What do Americans Think About Labor Unions & Strikes? 75% of Americans support workers going on strike  71% of Americans support labor unions  Support for Oppose There is a policy proposed to implement a national law to give all employees the right to collective bargaining and wage negotiations.
Labor organizing thought listing prompt	Please carefully consider whether or not you support this policy. Before you indicate your level of support or non-support, please tell us what goes through your mind, one thought at a time, when evaluating this policy. List any thoughts that come to mind about this topic. Please enter your thoughts one at a time below.
Carbon capture & storage policy context	We would like to know what you think about a policy that increases annual federal funding to the Department of Energy (DOE) for carbon capture and storage (CCS) programs. We will ask you how much you support the policy after you read about this technology. In many local communities across the country, there are CCS projects planned and under construction - yours may very well be

	one of them. Over 200 CCS projects are planned or under construction in the U.S. The Growth of Carbon Capture and Storage (CCS) in the U.S. Sources: Congressional Budget Office, Clean Air Task Force, Global CCS Institute For over a decade, the U.S. federal government has offered financial incentives to encourage the development and implementation of technologies designed to capture carbon dioxide (CO₂) emissions, aiming to lower the country's CO₂ output. Recent laws have greatly increased the yearly investment in these initiatives. Carbon capture and storage (CCS) is a process that captures carbon dioxide (CO₂) emissions from industrial sources, and then transports it to locations where it is securely stored beneath the Earth's surface. This method consists of three main stages: capture, transport, and storage, each incurring expenses for the businesses implementing this technology. One method of storage involves the injection of CO₂ into oil reservoirs to enhance oil recovery and can generate financial benefits. This practice is commonly adopted to offset some of the expenses associated with CCS. Currently, there are fifteen active CCS facilities in the United States, with a combined capability to capture 0.4% of the country's annual CO₂ emissions. There are also 214 CCS projects either being built or in the planning stages. Should these projects be successfully completed, they would boost the country's CCS capacity to 3 percent of its current annual CO₂ emissions.  CCS has been found to have regional health benefits, particularly for those with asthma. There are also risks associated with CCS plants; compressed carbon can leak at different parts in the process and has negative impacts on human/animal health risk and on the environment. Businesses that capture and store CO₂ qualify for the "section 45Q federal tax credit", which encourages the adoption of CCS and decreases federal income. The Treasury reported that from 2010 to 2019, companies received \$1 billion in such tax credits. The use of CCS is still rare: a probable explanation is that the expenses associated with its installation and operation often surpass any financial benefits companies might gain from using it.
Carbon capture & storage norm treatment	Next, you will view a short animation that visualizes distrust in the local governments that are responsible for regulating and permitting new CCS projects. 10,000 people answered: do you trust or distrust the local governments that are responsible for regulating and permitting new CCS projects? Each box represents the opinion of 100 individuals that were surveyed-- green means trust and red means distrust. Please note that this data is accurate and comes from researchers at Princeton University.

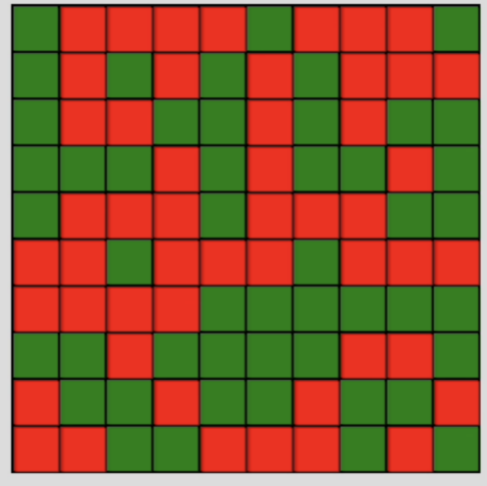
	 <i>Still capture from video animation.</i>
Carbon capture & storage thought lister prompt	Do YOU support or oppose a policy that increases annual federal funding to the Department of Energy (DOE) for CCS programs?

Table A4.6. Text and images used in stimuli for Study 2a-c.

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	-0.305	-0.534	-0.08	0.0086 **
ADE	-0.155	-0.339	0.02	0.0870 .
Total Effect	-0.460	-0.746	-0.17	0.0018 **
Prop. Mediated	0.663	0.290	1.08	0.0092 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 355

Simulations: 10000

Table A4.7. Study 2c mediation analysis. The mediating effect of proportion of support thoughts in CCS policy context is examined.

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.126079	0.004548	0.26	0.0410 *
ADE	0.181077	-0.000636	0.37	0.0510 .
Total Effect	0.307155	0.091774	0.53	0.0048 **
Prop. Mediated	0.410472	0.017465	0.98	0.0442 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 375

Simulations: 10000

Table A4.8. Study 2a-b SMRD mediation analysis in energy efficiency and labor protections policy context. Treatment is interacted with policy topic and controlled for political party.

Stimuli Section	Detail
Context/Control Stimuli	<i>Speaker 1 0:00</i> Green energy companies are spending big bucks on solar farms in rural areas. The building boom is raising concerns that short term profits for landowners could cause long term damage for communities. Mark Meredith is live in Richmond, Virginia with a closer look. Hi Mark. <i>Speaker 2 0:15</i> Dana, good morning to you. Well, there is a great debate underway here in Virginia and states all over the country about how much land should be devoted to these solar farms, and in some of the smaller communities, like the one we visited here in Virginia, the debate could be quite contentious. When we're talking about solar farms, we're essentially talking about large areas of land that have been cleared out to make way for hundreds, if not 1000s of different solar panels. For these farmers that lease the land, the power is then either transferred to nearby communities or sometimes added to the larger power grid. Energy companies offering a lot of money to these farmers, sometimes \$1,000 an acre, to lease their land for these projects.
Positive Norm Stimuli	<i>Speaker 1 0:00</i> Solar is the fastest growing renewable energy source, up nearly eight-fold in the last decade. One forecast shows the Inflation Reduction Act will lead to a 48% boost in solar deployment in the next 10 years. In the two years since the new rebates, Georgia Interfaith Power and Light says it has seen a 50% increase in faith communities enrolling in their solar-wise program.

	<i>Speaker 2 0:23</i> Solar is one of the most popular things in America right now. You know, 89% of Americans support solar. That's about 85% of Republicans, 95% of Democrats. Solar is creating jobs at 12 times the pace of the rest of the economy. So, you know it's one of the few bright spots. Universally popular. <i>Speaker 3 0:46</i> Climate change. It's been a political lightning rod. But according to the Yale program on Climate Change Communication at George Mason University, 78% of registered voters support providing tax rebates for folks who produce energy efficient vehicles or solar panels, while 79% say they support funding more R&D into renewable energy sources like solar and wind. And again, remember, this is super majority support at a time when our politics seem hopelessly divided.
Negative Norm Stimuli	<i>Speaker 1 0:00</i> A new report from Pew Research Center finds that although many Americans do want to reduce their carbon footprint, they're actually taking fewer steps to do so. The main part of this study focused on renewable energy, specifically solar and wind. Why are people supporting that, investing into that less than before? <i>Speaker 2 0:16</i> Well, a brand-new NBC News poll finds that only 5% of voters out of 100 think climate change is the most important issue facing this country right now. <i>Speaker 3 0:24</i> Alan Anglyn owns 40 acres of property in southwest Johnson County, Kansas. Now he's facing a massive solar field about 150 feet from his property line, taking away from its rural beauty. <i>Speaker 4 0:38</i> We bought out here, built out here, have our life savings invested in these places, and we see it as detrimental to, you know, character of the community. We see it as detrimental to our property values. <i>Speaker 3 0:49</i> One project in the works in Kansas would be massive, and residents are looking at two years of construction in a populated area and plenty of noise and dust. <i>Speaker 5 0:59</i>

	Fewer people are buying rooftop solar panels for their homes. The industry is shrinking, hit by high interest rates, tight credit and cuts to solar incentives. More than 100 solar contractors have already gone out of business during the past year, as demand has dried up.
--	---

Table A4.9. Transcripts from Study 4 video stimuli.

<i>Predictors</i>	value		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	3.60	3.04 – 4.16	<0.001
name [negative_num]	-1.18	-1.91 – -0.45	0.002
name [positive_num]	2.65	1.93 – 3.37	<0.001
Random Effects			
σ^2	1.29		
τ_{00} ResponseId	0.27		
ICC	0.17		
N ResponseId	20		
Observations	59		
Marginal R^2 / Conditional R^2	0.626 / 0.690		

Table A4.10. Study 4 pilot results to validate video stimuli. The table shows the results of a mixed effects linear regression; the perceived sentiment toward solar energy is regressed on the video type (policy context, positive norm information, negative norm information) with a random effect for participant ID. Sentiment toward solar energy is measured on a 7-point Likert scale from “very negative” to “very positive”. The reference category for video type is the neutral policy context video.

<i>Predictors</i>	Policy Support		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	3.35	3.01 – 3.68	<0.001
Support Norm Info.	-0.49	-1.00 – 0.01	0.055
Low Skepticism	-0.36	-0.78 – 0.06	0.089
Support Norm Info. x Low Skepticism	1.40	0.79 – 2.02	<0.001
Observations	378		
R ² / R ² adjusted	0.085 / 0.077		

Table A4.11. Linear regression where policy support is regressed on treatment and skepticism.

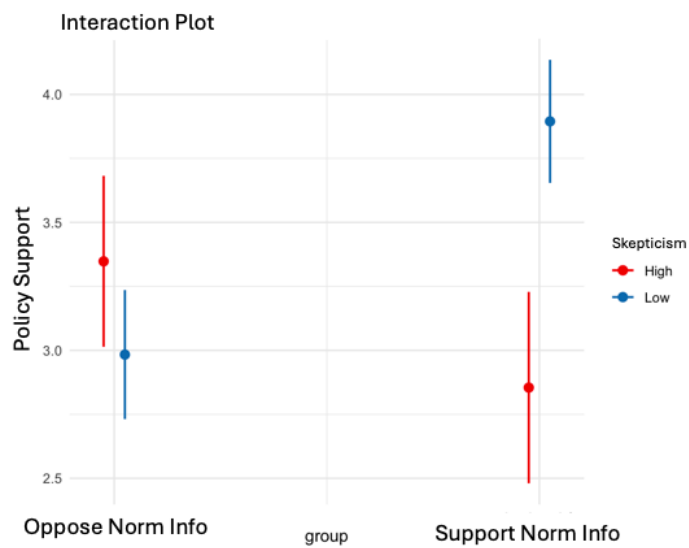


Figure A4.1. Interaction plot for regression shown in Table A4.11.

Causal Mediation Analysis				
Nonparametric Bootstrap Confidence Intervals with the Percentile Method				
	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.1820	0.0196	0.35	0.0284 *
ADE	0.2792	0.0440	0.52	0.0186 *
Total Effect	0.4613	0.1737	0.75	0.0028 **
Prop. Mediated	0.3947	0.0602	0.82	0.0304 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Sample Size Used: 378				
Simulations: 10000				

Table A4.12. Study 4 moderated mediation. This table shows the results of a moderated mediation analysis examining mediating effect of SMRD and moderating effect of skepticism on the relationship between norm treatment and policy support.

Causal Mediation Analysis				
Nonparametric Bootstrap Confidence Intervals with the Percentile Method				
	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.26511	0.05278	0.48	0.0132 *
ADE	0.19617	0.00237	0.39	0.0470 *
Total Effect	0.46128	0.17461	0.74	0.0014 **
Prop. Mediated	0.57473	0.19809	0.99	0.0138 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Sample Size Used: 378				
Simulations: 10000				

Table A4.13. Study 4 moderated mediation. This table shows the results of a moderated mediation analysis examining mediating effect of SMRD and moderating effect of skepticism on the relationship between norm treatment and policy support.



Figure A4.14. Policy support measured in Study 5 across the four process conditions. Policy support is measured on a 5-point Likert scale, the conditions are noted on the x-axis.