Three Essays on Information and Political Belief Change



Thesis submitted in partial fulfilment of the requirements for the DPhil in Politics

Andrew Lewis
Department of Politics and International Relations
University College, Oxford

Hilary, 2023

Abstract

How do our preconceptions and prejudices affect the way we evaluate political information and actors? This question motivates each of the articles that comprise this dissertation. In **One Fact, Two Beliefs (Article 1)**, I tackle head on the question of how responsive misperceptions are to disconfirming information, presenting a novel context in which to evaluate not only whether but also how people update their beliefs in response to new evidence. I find that, contrary to some renderings of modern political actors as fact-resistant, people do generally converge their beliefs to align with credible new information — yet, the lessons people draw from new information are systematically inclined to favour their preferred view, such that the effect may be to exacerbate rather than alleviate belief and attitude polarisation. In **Does Candidate Identity Affect Voter Response to a Scandal? (Article 2)**, I seek to answer the titular question, adapting and adding to an experimental toolkit for understanding the powerful effect of anti-prejudice norms (or lack thereof) on the suppression (or revelation) of prejudice. In so doing, I find evidence of a strong effect of prejudiced thought on evaluations of political actors. In **Give US the Ballot? (Article 3),** I address an information asymmetry arising from people's dramatically different voting experiences, and set out to understand views about voting impediments. I further test the receptiveness of these beliefs to new information, highlighting the effects of long waits on democratic participation, and find people across the political spectrum to be receptive. This translates to significantly increased support for policies that would make it easier for all Americans to participate in democracy, with none of the partisan differences that characterise other, previously polarised issues around voting access and elections.

Whatever colours you have in your mind,

I'll show them to you and you'll see them shine

Bob Dylan

Acknowledgements

With all of the good fortune I have had in my life, I could fill another dissertation with thanks. I would not have found this path without the guidance and encouragement of many people kinder and smarter than myself along the way.

For teaching me why, and showing me how, we might seek to understand behaviour, I am always indebted to George Loewenstein. For educating me on the world, and reminding me I only get one go in it, I am very grateful to Silvia Borzutzky. For inciting my interest in the topics herein, I send overdue thanks to Rodger Morrow, Melissa Willard-Foster, Shereen Chaudhry, and Lee Branstetter.

My time at Oxford has been truly enriching. I owe many thanks to my supervisor, Ray Duch, for the kindness he has shown and opportunity he has afforded me over the years. The work herein greatly benefitted from the thoughtful criticism and advice of Ray, Noah Bacine, James Tilley, Pepper Culpepper, Stephen Pettigrew, Inga Steinberg, Leonie Westhoff, Patrick Vu, and the many other researchers I've had the pleasure of working with at CESS and the DPIR. All of my ideas would be worse but for Alexia Tragakes. I am grateful to CESS for generous funding of these studies, and to the Royal Society for funding my research that did not make it into these pages, with special thanks to Areeq Chowdhury.

I dedicate this dissertation to my beloved mom and my late, great dad — the sources of any good qualities I may possess. To my big and incredible family and to my loving friends, I owe more than I could ever repay. I cherish you and thank you for always seeing the best in me.

Contents

List of Tables ..... **vi**
1 Introduction **1**
2 Literature Review **9**
2.1 Heuristics and (Political) Judgments 10
2.2 Beliefs and Behaviours 12
2.3 What's to Come 17
2.4 References for Chapters 1 & 2 17
3 One Fact, Two Beliefs: How Partisans Update Misperceptions **22**
3.1 Introduction 23
3.1.1 A Changing Climate? Changes and Gaps in Americans' Climate Beliefs 27
3.2 Literature Review 31
3.2.1 Motivated Thought 31
3.2.2 Motivated Political Judgment and Belief 34
3.2.3 When Beliefs Do (and Do Not) Change 37
3.2.4 Corrected Misperceptions and Their Consequences 40
3.2.5 Research Puzzle 42
3.3 Experimental Design & Methods 48
3.3.1 Study Design 48
3.3.2 Methods 49
3.4 Results 52
3.4.1 Existing Gap in Beliefs 52
3.4.2 Do Beliefs Converge? 53
3.4.3 Direction of Belief Change 57
3.4.4 Is There Backfire? 64
3.4 Discussion 66
3.5 References 70
A.3 Appendices 77
4 Does Candidate Identity Affect Voter Response to a Scandal? **83**
4.1 Introduction 84
4.2 Literature Review 87
4.2.1 Stereotypes, Prejudice, and Discrimination 87
4.2.2 Measures of Prejudice 90
4.2.3 Norms and Behaviour 92
4.2.4 Does Prejudice Affect Voting Behaviour? 94
4.2.5 Transgressions and the Activation of Prejudice 98

4.2.6 Research Puzzle: Prejudice as Political Belief ... 101
4.3 Political Scandals in the Wild ... 105
4.3.1 Constructing an Observational Dataset ... 105
4.3.2 Forming Expectations Based on Observational Data ... 106
4.4 Experimental Design & Methods ... 112
4.5 Results ... 115
4.5.1 Summary of Main Experimental Results ... 115
4.5.2 Examination of Main Experimental Results ... 117
4.5.3 Heterogenous Treatment Effects ... 123
4.6 Discussion ... 126
4.7 References ... 129
A.4 Appendices ... 135
5 Give US the Ballot? Dynamism in Beliefs on Voting Access ... 163
5.1 Introduction ... 164
5.2 Literature Review ... 167
5.2.1 The State of the Wait ... 167
5.2.2 Who Has to Wait? ... 168
5.2.3 Do Wait Times Matter? ... 170
5.2.4 Americans' Beliefs on Voting Access ... 172
5.2.5 Receptiveness of Beliefs to Information ... 173
5.2.6 Research Puzzle ... 175
5.3 Experimental Design & Methods ... 177
5.3.1 Design ... 177
5.3.2 Experimental Conditions ... 177
5.3.3 Dependent Variables ... 182
5.3.4 Covariates and Ancillary Questions ... 183
5.3.5 Analysis Plan ... 184
5.4 Results ... 185
5.4.1 Topline Survey Results & Past Voting Experiences ... 185
5.4.2 Information Treatment Versus Control ... 194
5.4.3 Identity-based Treatment Effects ... 198
5.5 Discussion ... 204
5.6 References ... 206
A.5 Appendices ... 211
6 Conclusion ... 220

List of Tables

Table 3.1: Mean favourability towards NG and MG by party 52
Table 3.2: Mean favourability towards NG and MG by condition 53
Table 3.3: Mean belief convergence by condition 54
Table 3.4: Effects of experimental treatments on belief convergence 54
Table 3.5: Effects of experimental conditions on belief convergence with covariates 55
Table 3.6: Effect of partisanship on downward NG adjustment 59
Table 3.7: Effect of partisanship on upward MG adjustment 59
Table 3.8: Effect of condition on downward NG adjustment 60
Table 3.9: Effect of condition on upward MG adjustment 61
Table 3.10: Downward NG adjustment with covariates 62
Table 3.11: Upward MG adjustment with covariates 63
Table 3.12: Extreme belief and backfire 65
Table 4.1: Mean difference between vote share in previous and subsequent election by party 107
Table 4.2: Difference in vote share with covariates 108
Table 4.3: Likelihood of re-election given running again 109
Table 4.4: Mean candidate favourability by condition 115
Table 4.5: Mean likelihood of voting for candidate by condition 116
Table 4.6: Linear regression for main DVs by condition (controls only) 118
Table 4.7: Linear regression for main DVs by condition (treatments only) 118
Table 4.8: Linear regression for main DVs, C3 vs C4, with covariates and condition x party interaction 120
Table 4.9: Linear regression for main DVs, T1 vs T3 121
Table 4.10: Tiered covariate regression models for favourability, T1 vs T3 122
Table 4.11: Mean candidate favourability by condition (Republican respondents) 124
Table 4.12: Mean likelihood of voting for candidate by condition (Republican respondents) 124
Table 5.1: Mean responses on primary outcome variables in C1, full sample 185
Table 5.2: Mean responses on outcome variables in C1, by partisan preference 186
Table 5.3: Effect of partisanship on outcome variables in C1 187
Table 5.4: Mean responses on trust in elections in C1 by partisan preference 189
Table 5.5: Effect of partisanship on electoral trust in C1 190
Table 5.6: Mean responses on legal and illegal voting in C1 by partisan preference 191
Table 5.7: Longest previous voting wait time, by race 192
Table 5.8: Longest previous voting wait time, by region 192
Table 5.9: Linear regression on DV responses, C1 vs T1 194
Table 5.10: Mean responses on support for reform, C1 vs T1, by partisan preference 195
Table 5.11: Covariate regression models on support for reform, C1 vs T1 195
Table 5.12: Covariate regression models on bad for democracy, C1 vs T1 197
Table 5.13: Linear regression on support for reform by condition (T1 vs. T2) 199
Table 5.14: Tiered covariate models on support for reform by condition (T1 vs. T2) 200
Table 5.15: Tiered covariate and interaction models on support for reform by condition (all treatments) 201

1

1 Introduction

How do our preconceptions and prejudices affect the way we evaluate new information? In one stylised telling of political and economic behaviour, our beliefs and actions follow an internally-consistent set of preferences, against which new information and opportunities are weighed and selected on the basis of their accordance thereto. In this view, derived from rational choice theory (RCT), our behaviours and decisions can thus be understood as efforts to satisfy our preferences — or, to maximise our *utility*. Such conceptions of behaviour are amongst the most influential contributions to the last century's study of political science (e.g., Arrow, 1951; Downs, 1957; Olson, 1965; Riker & Ordeshook, 1968; Fiorina, 1981; Aldrich, 1993). Following in the vein of economics, RCT introduced to the literature a formal theoretical framework by which practitioners could predict individual behaviour, enabling the explanation of socio-political phenomena, such as, for example, why people choose to vote.

From this perspective, though most voters are assumed to use *low-information* rationality to inform their views (Popkin, 1991), new information should nonetheless provide dynamism to our beliefs and actions. While voters lean on heuristics to make their voting choices, they keep a "running tally" (Fiorina, 1981, p.84) of political party evaluations, and modify their partisan preferences and voting behaviours on that basis. With clearly defined and complete preferences, the rational actor "proceeds toward its

goals with a minimal use of scarce resources and undertakes only those actions for which marginal return exceeds marginal cost" (Downs, 1957, p.137). As such, our primary task when faced with relevant information — to the extent that acquiring that information is costless — is to factor it into our beliefs and behaviours such that our higher-order preferences are satisfied. Thus, assuming individuals have a desire to hold accurate views, beliefs should be responsive to credible information irrespective of how it corresponds to previously held beliefs and attitudes.

Yet, there is a growing body of evidence[1] that voters, particularly in the US and the UK, are becoming more entrenched in their beliefs, and less responsive to countervailing evidence. Or in other words, ideological and other predispositions are altering the efficacy of new information, and how it is evaluated, based on its content. In turn, rather than preferences driving partisan identification, some posit the relationship to work in reverse — that is, partisanship as an "unmoved mover" (Johnston, 2006, p.329), staying fixed while pushing other beliefs and values in line with dominant party views. A logical consequence of unmoving partisan allegiances is increased polarisation, as voters on both ends of the spectrum more consistently align beliefs across a range of separate issues (Levendusky, 2010). To that end, constructing an index of nine widely-used measures, Boxell and co-authors (2017) find American polarisation to have worsened for every age cohort over the period 1996 to 2012, and particularly so for voters over 65. Notably, in line with the previous, they find a marked uptick in the likelihood that individuals' views across a range of issues track the

[1] Unsurprisingly, even the academic debate on polarisation is itself polarised. For example, both Fiorina and Abrams (2008) and Glaeser and Ward (2006) argue against the view that polarisation is on the rise — both deeming such a consensus a 'myth'. However, the weight of evidence in subsequent years has shifted against this perspective (see, for example, Levendusky, 2010; Westfall et al, 2015); most researchers accept that polarisation, and indeed affective polarisation (Iyengar et al, 2019), is deepening, and set out to investigate the causes and consequences thereof.

dominant views of their respective party (i.e., ideological consistency) — a result mirrored in surveys conducted by the American polling firm Pew (Pew, 2014).

Taking these literatures in concert, much scholarship is now focused on how individuals' *directional* motivations (e.g., partisan, ideological, identity-based) influence their propensity to update beliefs on the basis of new information (Kunda, 1990). The subtext beckons an important question: what is the relationship between information and beliefs, attitudes, and behaviours? Insights from political and social psychology detail the effects of our previously held beliefs on how we process and evaluate new information, suggesting that — rather than amending our beliefs on the basis of new information — the way we select, process, and evaluate new information is in fact often dependent on our previously held beliefs, such that we may choose to retain or discard information depending on how it corresponds thereto. This area of enquiry focused on *motivated cognition* — comprised of insights from the fields of political and social psychology, judgement and decision-making, and behavioural science — forms the basis of some modern models of political behaviour (for example, Lodge & Taber, 2000).

For two reasons, this notion of predisposition-dependent belief change is crucial to questions of political importance. First is the axiom that an informed populace is a precondition for a functioning democracy. Often attributed to Thomas Jefferson[2], this sentiment has been echoed in a variety of scholarship in both political theory and empirical political science. For example, in his seminal work on the social requisites of democracy, Lipset writes that education, if not a sufficient condition for democracy, "comes close to being a necessary condition in the modern world" (Lipset, 1959, p.80). The importance of this relationship was similarly underscored by the

[2] In a letter to Charles Yancey, dated 6 January, 1816, Jefferson wrote "if a nation expects to be ignorant [and] free, in a state of civilisation, it expects what never was [and] never will be." Quoted from the Monticello collections at the Library of Congress (retrieved at: http://tjrs.monticello.org/letter/327).

philosopher John Dewey, who wrote that: "[w]hat nutrition and reproduction are to physiological life, education is to social life" (Dewey, 1916, p.8).

Of course, belief polarisation does not necessarily imply a lack of education, nor the lack of an informed public. There are numerous issues (political and otherwise) on which it is entirely reasonable for people to disagree. For example, attitudes toward social issues (such as, say, the use of illicit drugs) tend to be more ideologically, rather than empirically, driven. However, there are also contexts in which views should coalesce. If the outcomes of interest (e.g., improved educational access) and the means of measuring progress toward that goal (e.g., percent of low-income students to receive financial assistance) are mutually agreed upon, all parties should interpret evidence corresponding thereto in a similar fashion. In such circumstances, if partisan (or other) predispositions supersede empirical reality, the result is a less-informed public, a lower return to the diffusion of information, and a weaker democracy.

In a similar vein, though research (reviewed in Article 2) documents a dramatic shift in prevailing social norms regarding the acceptability and prevalence of identity-based stereotypes and prejudice, evidence suggests such identity-driven prejudices, and resulting instances of discrimination, are still frequent in interpersonal behaviour. Much like how ideologically- or partisan-based predispositions affect the interpretation of new information, these prejudices have the potential to alter the way political actors are evaluated on the basis of identity. For example, a central understanding put forth by the psychological literature on motivated cognition is that individuals subject pro- and counter-attitudinal information to different evidentiary standards. That is, people will more readily accept information that agrees with their opinions, while exercising more scepticism toward opposing information (see, for example, Kunda, 1990; Ditto & Lopez, 1992). On this basis, if an individual holds a slight favourable prejudice (conscious or unconscious) toward a particular identity, he might be more sceptical of

negative information about a politician of that identity group. Conversely, he may more readily believe, or more heavily weight, negative information about a member of an out-group.

In turn, both types of predispositions have a material effect on political opinions, and these opinions form the basis of subsequent behaviours. This is the second key reason to focus on predispositions and information processing: to give context to and improve the accuracy of our study of behaviour. To examine political behaviour without simultaneously interrogating its perceptual and cognitive bases is to miss the interdependency there between. Ultimately, our perceptions (right or wrong) of economic, social, and political phenomena powerfully shape our attitudes toward them — and in turn our attitudes shape how we subsequently perceive new information and form and change beliefs.

The focus of this dissertation, then, is on this interplay between information and beliefs, and between beliefs and behaviour. Each of the three articles herein is focused on how dynamic political beliefs and behaviours are (or are not) to new information and contextual changes, and how a particular component of individual identity (e.g., partisanship) influences the evaluation of information. Taken together, the overarching theme that unites this area of research is how individuals evaluate and make use of new information to form and adjust their political opinions. The aim is to synthesise and expand upon insights from the fields of political, psychological, and behavioural science in order to advance our understanding of political beliefs, with a particular focus on how predispositions affect the efficacy of information.

The rest of this document proceeds as follows. In Chapter 2, I review a wide body of social and political psychological research on information processing, prejudicial and categorical thinking, and motivated belief and belief change, and how

these underlie the conceptualisations and evaluations of the research questions examined subsequently.

In **One Fact, Two Beliefs (Article 1)**, I approach directly the question of how responsive beliefs are to new, disconfirming information. Specifically, I address an oddity in people's beliefs about methane-based gases: significantly greater favourability theretoward when these gases are labelled as *natural gas* versus *methane gas*. This context provides a unique opportunity to consider belief change not only as it pertains to the simple question of whether people are likely to accept factual corrections, but also how their subsequent views reflect any such changes. In line with the understandings of factual receptivity and motivated information assimilation put forward in the article, I show that, contrary to some claims that people are fact-resistant, individuals seem to have no problem adjusting misperceptions to incorporate new, credible information. Yet, perhaps more importantly, their interpretations of what these facts actually mean differ systematically and predictably based on their broader partisan attitudes. The upshot is that, though increased information may bring people's beliefs more in line with evidence, in a polarised political context this is not only unlikely to bring partisans closer to agreement, but in fact may exacerbate existing divides in beliefs and attitudes.

In **Does Candidate Identity Affect Voter Response to a Scandal? (Article 2)** I develop a novel context in which to evaluate whether and how gender- or race-based prejudice might affect people's evaluations of political actors. I take as a starting point a rich body of past research demonstrating that people's overt expressions of prejudice, particularly in survey responses, conceal the true extent of held prejudices (conscious or unconscious), and the ramifications thereof when prejudice translates to discrimination. To examine this assertion and its potential consequences, I adapt a common experimental design in which participants evaluate identical individuals (in this case, hypothetical candidates for local office) with only the difference of the individuals'

names, which signal gender and race. Incorporating evidence regarding the powerful effect of social norms (particularly anti-prejudice and egalitarian norms) on expressions of prejudice, I innovate on this design by considering responses to these candidates under two normative contexts: the first, when the information provides only a simple candidate biography (controls), the second, when this information includes a brief article about a supposed ethical transgression the candidate has committed (treatments). The latter conditions proxy the effect of a political scandal, or other deviations from ethical ideals, on how political actors are evaluated. In turn, this allows for comparisons not only between candidates, but also based on the within-candidate variation in ratings in the presence or absence of a scandal. In line with recent research, I observe no evidence of a gender gap in people's evaluations of female politicians, with or without scandal. Conversely, I find strong evidence of racial discrimination, but only in the presence of a scandal: that is, comparing candidates between the controls, there are no significant differences in people's favourability toward or likelihood of voting for candidates (all of which are in the order of 70%) — but, when candidates are perceived to have transgressed, Black candidates and particularly Black male candidates are punished significantly more severely than their white counterparts. I interpret these findings in light of the literatures reviewed on how prejudices and preconceptions affect subsequent judgments, and how the presence of an ulterior justification to which to attribute one's prejudice greatly increases the likelihood of discrimination.

Finally, in **Give US the Ballot? (Article 3)**, I address an information asymmetry that results from people's divergent past experiences voting: namely, the dramatic and highly-concentrated effects of long wait times to vote on a small but meaningful segment of the electorate (roughly five million voters in a given national election cycle). Past research on the topic finds that these waits meaningfully affect people's decisions to or not to vote, in both present and subsequent elections, and that

these decisions bear on ultimate election outcomes. In the control condition, which serves as a resource for policymakers wishing to understand people's views about impediments to voting, I find that people across partisan divides generally favour policies aimed at curbing long queues. Moreover, despite some meaningful differences in people's trust in elections based on their partisanship, measures of trust for all parties indicate people are more trusting than untrusting of American elections. In the treatments, I find that these views are dynamic to new information, translating to significant increase in support for reform, and, crucially, that there are no partisan differences in the effect of information on belief change. I interpret the receptiveness of these views — but the lack thereof on related issues, such as voting by mail — in line with a broad body of past research demonstrating the ways in which previous considerations, subject to the forces of partisan polarisation and elite cues, undermine people's propensity for belief change.

Chapter 6 concludes with a review and interpretation of each of these findings, and what they signal, individually and collectively, in regards to the original question of how preconceptions and prejudices affect the evaluation of political issues and actors.

2

2 Literature Review

"If therefore those that pass for principles ... are only made to us by blind assent, we are liable to be misled by them; and instead of being guided into truth, we shall, by principles, be only confirmed in mistake and error" John Locke, *An Essay Concerning Human Understanding* (1690, bk.4, ch.12)

"If it's sometimes easier to reject strong evidence than to admit that we've been wrong, this is also information about ourselves worth having" – Carl Sagan, *The Demon-Haunted World* (1995)

It is hardly a new observation that the beliefs we hold affect the ones we are likely to come to hold. This notion, expressed in thousands of forms over perhaps as many years, is a truism of human reasoning and judgment, and is indeed reflected in common idioms such as *wearing rose coloured glasses* or *seeing what you want to see*. It is also the sum of a great body of research into psychology and cognition, which posits and provides evidence that people tend to favour and more heavily weight information that agrees with their pre-existing beliefs (Kunda, 1990; Ditto & Lopez, 1992), to seek evidence in support of those beliefs (Snyder & Swann, 1978), or to avoid information that potentially contradicts those beliefs (Golman, Hagmann & Loewenstein, 2017) — all at the cost of improved decision-making.[3] In turn, this biased assimilation of information has the dual effects of redoubling existing divides in beliefs and attitudes and thus exacerbating polarisation between disagreeing parties (Lord, Ross & Lepper, 1979).

While this view of how people process and evaluate information has seeped into the fabric of many social sciences, it remains a source of some debate amongst political science scholars. In the two prominent, and conflicting, theories of political behaviour outlined in Chapter 1, people are either rational, weighing any and all

[3] In this case *improved* is descriptive rather than normative, as these patterns of thought have been shown to yield decisions that stray from people's previously-stated preferences.

(costless) relevant information that may improve the accuracy of their evaluations and updating their behaviours in line with it; or they are motivated actors, selectively seeking (Sharot & Sunstein, 2020), injudiciously processing (Taber, Cann & Kucsova, 2009), and insufficiently incorporating (Jerit & Barabas, 2012) evidence into their evaluations so as to maintain consistency with a preferred belief or worldview. This dissertation draws in large part from the latter framework, critically reviewing and contributing to literatures on belief change and maintenance (Article 1), prejudices and discrimination (Article 2), and receptiveness to new information (Articles 1 and 3). The present section briefly surveys some of the common literature that will be encountered in the chapters that follow, and gives insight into how it will bear on the research questions that are presented and evaluated.

2.1 Heuristics and (Political) Judgments

"The human mind must think with the aid of categories" (Allport, 1958, p.19). To begin at the root, I first consider the nature of (political) judgments and decisions. As Allport wrote, people are categorical thinkers, relying on a number of shorthand rules and intuitions that help simplify and guide us through the complexities of our surroundings. For the most part, these heuristics are quite efficient and adaptive, helping us preserve scarce cognitive resources and yielding judgments that are no worse than those made under conscious scrutiny (Tversky & Kahneman, 1974[4]; Gigerenzer, 2008). In some common sketches of cognition, human thought and behaviour is the result of two distinct processes: the automatic processes, entailing heuristics and categorical thoughts, and the *controlled* processes, operating more consciously and applying reasoned consideration (Shiffrin & Schneider, 1977).

[4] Though often cast as pioneers of the so-called *negative heuristics* movement (perhaps rightfully given the trickle-down effects of their findings), Tversky and Kahneman were under no illusion that heuristics were *bad* when they set out to document some shortcomings of automatic thinking — noting in the abstract of their famous 1974 paper *Judgment Under Uncertainty* that heuristics "are highly economical and usually effective" (Tversky & Kahneman, 1974, p.1).

Sometimes termed *System 1* and *System 2*, these processes complement one another, allowing us to turn our attention only to matters sufficiently meriting the effort (Stanovich & West, 2000; Kahneman & Frederick, 2002). Yet, they can also be contrasting, yielding different evaluations of the same subject-matter depending which mode of thought is employed. In particular, the improper adaptation or application of heuristics yields a number of predictable decision errors, sometimes referred to broadly as biases (see, for example, Tversky & Kahneman, 1974). This can occur when a heuristic is based on flawed reasoning (e.g., in making incorrect attributions based on a misunderstanding of causality; Gilovich & Savitsky, 2002), or when a well-adapted heuristic is incorrectly applied (e.g., an American looking to their left before crossing a London street). Thus, we rely on heuristics and implicit associations to simplify our judgments and beliefs, but we also rely on the controlled processes to override the unconscious processes when they may be leading us into traffic.

In various conceptualisations, the automatic processes of System 1 are thought to entail either cognitive or affective cues, or sometimes both. In the view that they are affective, automatic processes are charged with emotion, and thus engender evaluations and behaviours that are more driven by positive or negative cues for how we should react to a given stimulus (whereas a cognitive heuristic need not hold any valence judgments) (Metcalfe and Mischel, 2000). These cues and related behaviours are thought to work proportionately, such that the strength of the affective cue is predictive of its impact on judgments and behaviours (Loewenstein, O'Donoghue & Bhatia, 2015).

The idea that human behaviour results from (at least) two separate processes has been echoed in "writings of the earliest philosophers to the present" (Loewenstein, O'Donoghue & Bhatia, 2015, p.55). In his *Essay Concerning Human Understanding*, the philosopher John Locke wrote of the "modes of thought" underlying belief formation,

as sensation, remembrance, and contemplation — with much similarity to the conceptualisations described above:

> "the perception or thought which actually accompanies, and is annexed to, any impression on the body, made by an external object, being distinct from all other modifications of thinking, furnishes the mind with a distinct idea, which we call SENSATION;—which is, as it were, the actual entrance of any idea into the understanding by the senses. The same idea, when it again recurs without the operation of the like object on the external sensory, is REMEMBRANCE" (Locke, 1690, bk.2, ch.9).

In the view of the theories covered already, the remembrance of sensation would bring with it either affective or cognitive heuristic associations, which would be applied in evaluation. As for conscious thought, Locke contends that contemplation has only occurred if an idea is held long enough to undergo "attentive consideration" (Ibid). While there is hardly a full understanding of when people will engage the more effortful conscious processes, much research has sought to elucidate some of the conditions that give rise to *attentive consideration*, with some arguing they are only activated in cases where automatic processes fail to form a satisfactory evaluation (Lieberman, Schreiber & Oschner, 2003).

2.2 Beliefs and Behaviours

The upshot of these underlying processes for the present research is the ways in which they bear on actual beliefs and behaviours. In a famous theory of why people maintain their beliefs, Kunda (1990) elaborates two goals that people weigh in making decisions: *directional goals*, or the desire to arrive at a certain conclusion, and *accuracy goals*, or the desire to be correct[5]. In her view, these are balanced thusly:

> "The biasing role of goals is thus constrained by one's ability to construct a justification for the desired conclusion: People will come to believe what they want to believe only to the extent that reason permits. Often they will be forced to acknowledge and accept undesirable conclusions, as they appear to when confronted with strong arguments for

[5] That is, with respect to the best available evidence.

> undesired or counter-attitudinal positions" (Kunda, 1990, p.483).

In this understanding, it is therefore the case that we follow directional goals to the extent that we can self-justify doing so. This means that, all else equal, we are more likely to incorporate belief-confirming information into our evaluations, and to discard that with which we disagree. Or, in another famous rendering, people subject pro-attitudinal information to a standard of *can I believe this*, while they subject counter-attitudinal information to a standard of *must I believe this* (Dawson, Gilovich & Regan, 2002). This model of *motivated scepticism* (Ditto & Lopez, 1992) essentially affects the amount and quality of information we need to arrive at a particular conclusion: needing less for things we are inclined to believe, and more for things we are resistant to believing. This model has been adapted into political psychology research to great effect.

For example, based on this past research, some scholars have argued that "all reasoning is motivated" (Lodge & Taber, 2000, p.186), and there is no good justification for thinking political judgments are excepted from this rule. Adapting the previous model of affective cues and so-called *hot* cognition, the authors contend that these decision cues, based on affective judgments of our past experiences and evaluations, bear strongly on how people consider political information and actors. In a series of experiments, they demonstrate the ways in which familiar political topics, when broached, are "charged with positive or negative affect" (Burdein, Taber & Lodge, 2006, p.365), which in turn dictates how we will react, evaluate, and behave. As I will argue in Article 3, this means relatively less-polarised political and policy areas are the most fertile grounds for belief change and consensus-building. In a particularly notable case, Americans' views on voting by mail — once a rather ancillary, administrative concern — became dramatically more polarised and prominent in the

wake of the 2020 election, wherein some well-known partisan actors levied (unfounded) claims that mail voting was susceptible to (and indeed, a primary means of) election rigging and vote fraud (Lockhart et al, 2020). These allegations of malfeasance — when levied by either political side — have serious consequences for people's trust in the functioning of elections and, by extension, faith in democracy itself (Albertson & Guiler, 2020; Clayton et al, 2021).

Beyond affective and cognitive cues, others have attributed biased political reasoning to people's deliberative capacities. In this view, it is argued, sophisticated partisans are more susceptible to partisan motivated reasoning, as they are better able to construct sound-seeming counterarguments (Kahan et al, 2012; but see Bago, Rand & Pennycook, 2020; Persson et al, 2021). Once again this returns us to Kunda's paradigm that we will follow our directional goals whenever we can self-justify doing so. Regardless of the mechanism, this *partisan perceptual bias* has been shown to affect not only evaluations of political issues and actors, but also knowledge thereof, such that we know more about topics for which the justifiable evaluations align more with our worldview, and less about those that may engender uncomfortable conflict — an effect exacerbated by the relative polarisation and salience of a subject (Jerit & Barabas, 2012). As I will cover in Article 1, these predispositions toward belief-consistent interpretation lead partisans to asymmetrically evaluate political topics ranging from how well the economy is doing (Duch, Palmer & Anderson, 2000; Tilley & Hobolt, 2011) to whether climate change is caused by human activity (McCright & Dunlap, 2011; Bolsen, Druckman & Cook, 2015). On the latter issue, politicisation is a particularly potent shaper of beliefs, with partisan cues significantly eroding propensity to adapt beliefs about global warming in line with evidence (Unsworth & Fielding, 2014; Benegal & Scruggs, 2018). Part of this comes down to what we deem credible evidence, with a significant partisan influence on these determinations (Druckman & McGrath, 2019).

Moreover, even if partisans could agree on credibility, people still tend to eschew uncomfortable information "to, in effect, license them to behave as they would really like to behave" (Golman, Hagmann & Loewenstein, 2017, p.129). Drawing from the previous section, it is thought that these effects result from a heuristic process sometimes dubbed *confirmatory hypothesis testing*, which stipulates that people approach a subject with a preferred conclusion in mind and find evidence to support it (Klayman & Ha, 1987). Thus, overly partisan beliefs can result from injudicious attention to, consideration of, and evaluations regarding pro- and counter-attitudinal information, and a general resistance to belief change (Stanley et al, 2020). Put simply, we need less (and less credible) evidence to maintain a current belief, and more (and more credible) evidence to change it.

One area where heuristics are particularly potent is in how we form beliefs about individuals based on characteristics we deem them as holding. Or, in common parlance, stereotypes and prejudices. In one explication, stereotypes are the cognitive shortcuts we use to evaluate people, while prejudices are the more affective cues that drive our behaviour (Fiske, 1998). As explained above, stereotypes — like all heuristics — need not be based on any valence judgments of stimuli as *good* or *bad*, whereas prejudices are definitionally judgments of valence (there is no such thing as neutral prejudice). A positive prejudice could be, for example, the view that Californians are particularly active and physically fit, which, though may perhaps lead to incorrect assumptions about an individual, is likely not particularly harmful. Much of the focus of prejudice research, then, is on *negative* prejudices, which can be thought of as "an unfair negative attitude toward a social group or a person perceived to be a member of that group" (Dovidio, 2001, p.829). The behavioural component here is discrimination, such that negative prejudices which are applied toward individuals are definitionally instances of discrimination (Fiske, 1998). As I review in Article 2, though the last century has

seen a dramatic shift in cultural and social norms around prejudice (and resulting instances of discrimination), research suggests that these forces remain latent in society, either actively — as in the case of people who openly hold prejudice — or unconsciously, as for those who desire not to be prejudiced but are nonetheless categorical thinkers with ingrained knowledge of common stereotypes. Nor is knowledge of stereotypes something people necessarily acquire by choice, but rather is acquired through socialisation and cues we pick up from people around us — or what has been called "culture in mind" (Hinton, 2017, p.2). This in turn can lead even those of us who very much wish not to be prejudiced to nonetheless apply stereotypes in our judgments (Lai et al, 2016). Once again, we fall prey to these heuristics because of the simple nature of our cognition, as we extrapolate from limited past experiences to form expectations of what current and future events may mean (Gilbert & Wilson, 2007).

How we moderate these forces is again a subject of debate. In one prominent model, people suppress prejudices for fear of either self- or social sanction — the former applying mostly to people who are not overtly prejudiced and thus hold relatively egalitarian views of all groups (Kunda & Spencer, 2003), the latter to those who are prejudiced but also conscious to the repercussions of violating social norms around intergroup relations (Crandall & Eshleman, 2003). The upshot of this research is that evaluating *expressions* of prejudice is an insufficient way of measuring its prevalence or absence: these forces are also latent, and depending on contextual and normative constraints, are more or less likely to factor into our behaviours. As will be seen in Article 2, I find strong evidence that there are significant differences in people's propensity to engage in race-based discrimination against hypothetical political actors based on these factors, and that prejudice reveals itself in the absence of adequate checks against it.

2.3 What's to Come

In the papers that follow, I find that these heuristics and resulting biases are present in consequential political evaluations and behaviours. They affect how we consider and make sense of new information, how we change our beliefs, and how we form perceptions of political events and actors. I also find that they can be overcome: that people are receptive to information when it seems to sufficiently justify belief change (Articles 1 and 3), particularly on matters that are not already the subject of polarised debate, and thus more charged with past belief (Article 3); and that social norms, or the external pressures that shape and delimit our behaviours, are powerful safeguards against the applications of heuristics and prejudices (Article 2). Nonetheless, we still fall prey to our categorical minds, interpreting evidence in a way that suits our preferences (Article 1), or acting on ingrained or believed stereotypes and prejudices when normative conditions or deliberative processes fail to override such judgments (Articles 2 and 3). In expounding on these findings in Chapter 6, I will argue that political research failing to account for these factors will also fail to understand and thus predict the forces that shape our beliefs and behaviour.

2.4 References for Chapters 1 & 2

Albertson, B., & Guiler, K. (2020). Conspiracy theories, election rigging, and support for democratic norms. *Research & Politics*, *7*(3), 2053168020959859.

Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science*, 246-278.

Allport, G. W. (1958). The nature of prejudice (abridged). Garden City, NY: Doubleday. (Original work published 1954)

Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica: Journal of the Econometric Society*, 404-437.

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of experimental psychology: general*, *149*(8), 1608.

Benegal, S. D., & Scruggs, L. A. (2018). Correcting misinformation about climate change: The impact of partisanship in an experimental setting. *Climatic change*, *148*(1-2), 61-80.

Bolsen, T., Druckman, J. N., & Cook, F. L. (2015). Citizens', scientists', and policy advisors' beliefs about global warming. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 271-295.

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). *Is the internet causing political polarization? Evidence from demographics* (No. w23258). National Bureau of Economic Research.

Burdein, I., Lodge, M., & Taber, C. (2006). Experiments on the automaticity of political beliefs and attitudes. *Political Psychology*, *27*(3), 359-371.

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, *129*(3), 414.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated Reasoning and Performance on the was on Selection Task. *Personality and Social Psychology Bulletin, 28*(10), 1379-1387.

Dewey, J. (1966). Democracy and education (1916). *Jo Ann Boydston (ed.). The Middle Works of John Dewey*, *9*, 1899-1924.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*(4), 568.

Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. Journal of Social Issues, 57(4), 829-849.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, *65*(2), 135-150.

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111-119.

Duch, R. M., Palmer, H. D., & Anderson, C. J. (2000). Heterogeneity in perceptions of national economic conditions. *American Journal of Political Science, 44*(4), 635-652.

Duch, R. M., & Stevenson, R. T. (2011). Context and economic expectations: when do voters get it right?. *British Journal of Political Science*, *41*(1), 1-31.

Fiorina, M. P. (1981). *Retrospective voting in American elections*. New Haven, CT: Yale University Press.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563-588.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 357–411). McGraw-Hill.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*(1), 20-29.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351-1354.

Gilovich, T., & Savitsky, K. (2002). Like goes with like: The role of representativeness in erroneous and pseudo-scientific beliefs. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 617–624). Cambridge University Press.

Glaeser, E. L., & Ward, B. A. (2006). Myths and realities of American political geography. *Journal of Economic Perspectives*, *20*(2), 119-144.

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. *Journal of economic literature*, *55*(1), 96-135.

Hinton, P. (2017). Implicit stereotypes and the predictive brain: cognition and culture in "biased" person perception. *Palgrave Communications*, *3*(1), 1-9.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., ... & Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, *22*, 129-146.

Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *The Journal of Politics, 74*(3), 672-684.

Johnston, R. (2006). Party Identification: Unmoved Mover or Sum of Preferences?. *Annual Review of Political Science*, *9*, 329-351.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732-735.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.*), Heuristics & biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review*, *94*(2), 211.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Lieberman, M. D., Schreiber, D., & Ochsner, K. N. (2003). Is political cognition like riding a bicycle? How cognitive neuroscience can inform research on political thinking. *Political Psychology*, *24*(4), 681-704.

Levendusky, M. S. (2010). Clearer cues, more consistent voters: A benefit of elite polarization. *Political Behavior*, *32*, 111-131.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy1. *American Political Science Review*, *53*(1), 69-105.

Locke, J. (1690). *An Essay Concerning Human Understanding: 25th Ed., In One Volume* (Book IV). Printed in 1825 for Thomas Tegg, 74, Cheapside; R. Milliken, Dublin; Griffin and Co., Glasgow; and M. Baudry, Paris. https://www.gutenberg.org/files/10615/10615-h/10615-h.htm

Lockhart, M., Hill, S. J., Merolla, J., Romero, M., & Kousser, T. (2020). America's electorate is increasingly polarized along partisan lines about voting by mail during the COVID-19 crisis. *Proceedings of the National Academy of Sciences*, *117*(40), 24640-24642.

Lodge M., & Taber C. (2000). Three steps toward a theory of motivated political reasoning. In A. Lupia, M. D. McCubbins & S. L. Popkin (Eds.), *Elements of Reason: Cognition, Choice, and the Bounds of Rationality* (pp. 183–213). Cambridge; New York: Cambridge Univ. Press.

Loewenstein, G., O'Donoghue, T., & Bhatia, S. (2015). Modeling the interplay between affect and deliberation. *Decision*, *2*(2), 55-81.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098.

McCright, A. M., & Dunlap, R. E. (2011). The politicization of climate change and polarization in the American public's views of global warming, 2001–2010. *The Sociological Quarterly*, *52*(2), 155-194.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological Review*, *106*(1), 3.

Olson, M. (1965). *The logic of collective action : Public goods and the theory of groups* (Harvard economic studies ; v. 124). Cambridge, Mass: Harvard University Press.

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*, 104768.

Pew Research Center. (2014, June 12). *Political Polarization and Growing Ideological Consistency*. Pew Research Center - U.S. Politics & Policy. https://www.pewresearch.org/politics/2014/06/12/section-1-growing-ideological-consistency/

Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. University of Chicago Press.

Riker, W. H., & Ordeshook, P. C. (1968). A Theory of the Calculus of Voting. *American Political Science Review*, *62*(1), 25-42.

Sagan, C. (2011). *The demon-haunted world: Science as a candle in the dark*. Ballantine books.

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, *4*(1), 14-19.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*, 127190. http://dx.doi.org/10.1037/0033-295X.84.2 .127

Stanley, M. L., Henne, P., Yang, B. W., & De Brigard, F. (2020). Resistance to position change, motivated reasoning, and polarization. *Political Behavior*, *42*, 891-913.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(5), 701-717.

Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior, 31*(2), 137-155.

Tilley, J., & Hobolt, S. B. (2011). Is the government to blame? An experimental test of how partisanship shapes perceptions of performance and responsibility. *The Journal of Politics*, *73*(2), 316-330.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124-1131.

Unsworth, K. L., & Fielding, K. S. (2014). It's political: How the salience of one's political identity changes climate change beliefs and policy support. *Global Environmental Change, 27*, 131-137.

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145-158.

3

3 One Fact, Two Beliefs: How Partisans Update Misperceptions

How responsive are incorrect beliefs to new information? Political science research offers varied perspectives on whether and how individuals update when presented with credible evidence a belief they hold is untrue. One difficulty in estimating propensity for evidence-based belief updating is the frequent comingling of facts and opinions, such that people may doubt how verifiable an assertion is (Walter & Salovich, 2021). The present research contributes to this literature by focusing on correction of a demonstrably false perception and measuring both the extent and direction of subsequent belief change. To do so, I exploit a unique oddity in Americans' beliefs: that is, greater favourability toward methane-based gases when framed as *natural gas* versus *methane gas*. Prior research suggests this perception gap is large and bipartisan: while roughly 76% of Americans report favourable attitudes towards *natural gas*, this figure is only 14% for *methane gas* (Lacroix et al, 2021). In reality, these terms refer to the same energy source, as methane is the primary constituent of so-called natural gases (Faramawy et al, 2016). To test the receptiveness of beliefs, I design and administer a survey experiment (n = 633) wherein participants indicate initial favourability towards both gases, are provided with a description of their likeness, and then are given a chance to (or not to) adjust their ratings. On the basis of past research, subsequent conditions test for any mediating effects of source credibility, partisan cues, and broader issue-attitudes on updating behaviour. The data show pre-treatment perception gaps in line with past research. I find individuals across conditions generally update their beliefs in the direction of the information, with no partisan differences in the extent of belief convergence. Conversely, there are significant differences in the direction of the individual belief changes causing this convergence, such that Democrats are more likely to adjust their natural gas beliefs downwards and Republicans to become more favourable toward methane gas. Updating behaviour is also sensitive to framing effects, with the addition of partisan cues significantly eroding propensity for belief convergence. I discuss these results and their implications for research on belief change and polarisation.

3.1 Introduction

Does increased access to information translate to a better informed public? To take a cue from accounts of political science that rely on the 'rational' model of behaviour (e.g., Downs, 1957; Fiorina, 1981), the answer would have to be yes. In such a model, individuals acquire and make use of all available costless information in order to make choices that best align with their set of preferences. Moreover, these preferences are sensitive to the content of that information, such that a change in evidence for or against some policy or action would translate to a realignment of an individual's preferences. Thus, assuming a desire for accuracy, high-quality information should be the determining factor on which individuals' beliefs are formed and adapted, and in turn these beliefs should shape their political and policy preferences.

Yet, a growing strain of literature in political behaviour documents how our prior beliefs and attitudes affect the way we seek out, consider, and ultimately make use of new information. This area of research rests on psychological accounts of motivated reasoning and belief maintenance, which demonstrate the ways in which people systematically prioritise belief- and attitude-consistent information over information with which they disagree (Ditto & Lopez, 1982; Kunda, 1990). Evidence from this field has created something of a new received wisdom: individuals' motivation to believe or not believe certain evidence is often the determining factor for how it is evaluated (Taber & Lodge, 2006).

In the context of deepening partisan polarisation in the US (see, for example, Levendusky, 2009; Westfall et al, 2015; Iyengar et al, 2019), this debate has taken on an increased urgency. Whereas the rational model views partisan attachment as a "running tally" of evaluations of a party's performance (Fiorina, 1981, p.84), the motivated model sees partisanship as an "unmoved mover" (Johnston, 2006, p.329), remaining fixed while driving other political beliefs — such as policy preferences or candidate

evaluations — to align with the dominant partisan dogma (Bartels, 2002). According to the former view, partisanship would not affect the extent to which people believe factual information, even when that information controverts their broader attitudes; according to the latter, partisan predispositions are the filter through which information is considered, and in turn affect policy opinions, political values, and voting behaviours. Summarising research on both sides, Johnston (2006, p.329) calls evidence for the partisanship-to-belief framework "remarkably robust."

Even as some research on political behaviour may be shifting toward the social-psychological model, there is hardly agreement on the extent to which people's political beliefs and attitudes are open to changing based on new evidence. In one telling, context is key, such that beliefs are less responsive to information on politically important and partisan polarised issues (Gaines et al, 2007; Druckman, Peterson & Slothuus, 2013). Conversely, other research finds that issue importance and salience can make individuals more likely to ground their views in evidence (Ciuk & Yost, 2016); that specific issue priors are more predictive than partisan priors in attentiveness to new information (Mullinix, 2016); and that political beliefs and even broader attitudes are responsive to new information (Tappin, Berinsky & Rand, 2021). A throughline in these somewhat conflicting findings is that partisans may acknowledge the same fact, but interpret its implications differently, such that the takeaway aligns with their preferred view (Tilley & Hobolt, 2011; Bisgaard, 2015).

In parallel to these debates about the role of credible information in belief maintenance and change, a growing body of evidence suggests that public misperceptions regarding important political and policy issues are widespread. For example, Flynn (2016) finds that one in five people confidently hold misperceptions about issues ranging from the size of the US foreign aid budget to gun control regulations (Flynn, 2016). On the latter issue, policy outcomes may be linked to these

misperceptions: some 77% percent indicate favourability toward federally mandated universal background checks on all gun purchases, but 41% erroneously believe such a mandate is already the law[6], potentially undermining demand for policy change (Aronow & Miller, 2016). Indeed, research even suggests misperceptions about the extent of polarisation itself, with people significantly overestimating the degree to which their out-party peers disagree on policy and social issues (Levendusky & Malhotra, 2016). One key context for research on misperceptions has been beliefs about climate change, where research (detailed below) finds large gaps in people's views and their responsiveness to new information.

For this reason, understanding how responsive people's beliefs are to new evidence is crucial to realising a better-informed public. In recent years, a host of research (reviewed below) has sought to answer this question, addressing misperceptions and information gaps and evaluating the factors that most influence people's propensity to update their views. These contributions have greatly advanced our understanding of when, why, and how political beliefs change. Yet, one limitation of much past research is the difficulty of separating facts from opinions (Walter & Salovich, 2021), potentially leading to an underestimation of people's openness to accepting new facts where they previously held a misconception. The contribution of the present research is to address a widely-held and demonstrably false belief, and to observe whether and how people update that belief under different conditions shown to affect likelihood of change.

To do this, I focus on the unique case of a fairly ubiquitous gap in Americans' favourability toward a fossil fuel energy source when it is framed as either *natural gas* (henceforth, NG) or *methane gas* (henceforth, MG). Prior research shows that, while a

[6] According to the American Academy of Pediatrics, federal background checks are only required for firearms purchased through federally-licensed dealers, which account for less than half of all gun sales in the US.

majority of Americans hold a generally positive view of NG, only 14% feel the same about MG, with a similar downward trend between both Republicans and Democrats (Lacroix et al, 2021). Yet, as the authors note, these terms are used to refer to the same energy source, as methane constitutes between 70-95% of NG, and indeed, pipeline grade NG used for heating and electricity is 95-98% methane (EPA, 2016).

In the experiment that follows, participants indicate their initial favourability toward NG and MG, and then are treated with information reflecting the above. Following the information treatment, participants are given the option to change their favourability toward both gases in light of the new information (or, of course, to maintain their initial views), as well as to answer brief questions about their broader climate and energy policy beliefs. Information treatments take four forms: the first (T1) is the simple presentation of this information, including a quote from the US Department of Energy explaining that NG "is primarily methane with small amounts of other hydrocarbons."[7] The second treatment (T2) presents the same information but removes reference to the source thereof, in order to test for how perceived credibility may alter the effect of information. The final two treatments present the same information as T1, but add additional partisan (T3) and issue-attitude (T4) cues, both shown to impact on updating behaviour in past research. These treatments are explained in more detail in the experimental design section that follows. The main outcomes of interest are whether beliefs converge (that is, whether individuals align their views on NG and MG) and how the underlying beliefs change (e.g., less favourable toward NG) — then, crucially, if and how this differs based on the framing posed in each condition.

[7] Link to the Department of Energy article here.

For three reasons, this context is particularly well-suited to examining belief change: first, the extent of misperception is similar between parties; second, the corrections pertain to an incontrovertible, physical fact; third, the presence of two variables enables the examination of not only the extent to which people align their beliefs, but also the direction of any belief change (e.g., whether NG ratings go down or MG ratings go up). The latter component gives a view into how these changes correspond to broader partisan and climate attitudes, and whether corrected misperceptions are likely to alleviate or exacerbate polarisation.

The rest of this paper proceeds as follows. Section 1.1 gives a background on Americans' views about climate change and fossil fuel energy sources. Section 2 covers literatures on motivated thought and political beliefs, and critically reviews a wide body of past research on corrected misperceptions and their consequences. Section 3 details the experimental design and empirical methods to be used in Section 4, which presents results. Section 5 concludes with a discussion of the findings and avenues for future research.

3.1.1 A Changing Climate? Changes and Gaps in Americans' Climate Beliefs

If global warming poses the greatest existential threat to society — as one recent United Nations (UN) report claims (UN, 2021) — one may not know it based on Americans' beliefs and behaviours. Recent surveys of public opinion in the US find roughly a third of all respondents believe climate change is "a small problem" or "not a problem at all" (Pew Research Center, 2021). Similarly, amongst a set of 31 issues, only 18% included climate change in their top three perceived problems facing the country, with inflation, the economy, crime and violence, and healthcare all rated as more important (Schulman et al, 2022). Americans are not alone in these beliefs. In a similar survey, only 16% of Britons mentioned climate change amongst their most important issues facing the UK,

down substantially from a peak of 40% in 2021 following the COP26 Climate Change Summit (Clemence, 2022).

Of course, important issues need not be mutually exclusive, and there are understandable reasons why, for example, people might worry more about inflation making their groceries unaffordable than the less proximal effects of climate change. Moreover, the last decade has seen considerable changes in people's climate beliefs, with two-thirds reportedly worrying about climate change in 2021, a roughly 27% increase over the same figure a decade prior (Saad, 2021). Yet, in spite of these changes, there remain important points of disagreement and cleavages between groups. For example, whereas 91% of Democrats fall into that category of worry, only 32% of Republicans expressed such concern (Saad, 2021).

Beyond the existence and threat of climate change, there are even greater divides in people's beliefs about the causes thereof, and the scientific consensus in support of those causes. While 88% of Democrats believe that climate change is caused by human activity (i.e., anthropogenic global warming, AGW), only 35% of Republicans share this view (Independents sit in the middle at 68%) (Saad, 2021). This further extends to the views of scientists, with a roughly 31% drop in the likelihood a Republican scientist believes in AGW compared to her Democratic peers (Bolsen, Druckman & Cook, 2015). As for consensus, while a meta-review of research on experts' beliefs finds that roughly 97% of all climate scientists endorse AGW (Cook et al, 2016), only one in ten respondents correctly estimate this level of consensus as greater than 90% (Leiserowitz et al, 2015). Once again these phenomena are not uniquely American: a survey of four European countries find that only about a third correctly estimate climate consensus as 'high' (i.e., greater than 80%) (Steentjes et al, 2017). Perhaps more troublingly, roughly a quarter of respondents placed the scientific consensus at 50% or lower.

In turn, views about the existence, causes, and consequences of AGW all factor heavily into policy preferences for mitigation thereof. For example, while the vast majority of Democrats want the US to reduce its reliance on fossil fuel energy sources, only about a third of Republicans agree (McCarthy, 2019). There are divides within parties as well, with two-thirds of Republicans who identify as conservative opposing efforts to transition toward carbon neutrality by 2050, while the inverse is true of those who identify as more moderate (Tyson, Funk & Kennedy, 2022). Past research suggests at least part of the reason for these gulfs in opinion stem from differing views about what people consider credible evidence of AGW (Druckman & McGrath, 2019). Such an explanation makes sense: because most of us are ill-equipped to parse the complexities of climate science ourselves, we necessarily rely on trusted sources to inform our beliefs. The Nobel Laureate Daniel Kahneman well-summarised this point in a recent interview, saying "I mean, I believe in climate change, but I have no idea about it really. What I believe in is the institutions and methods of people who tell me there is climate change."[8] Or, as the sociologist Nikolas Luhmann described it in his book *Trust and Power*, in a highly atomised and specialised society like our own, trust is essential to belief. To borrow his example, rather than learning how to fix our own car engines, we simply learn *who can* fix those engines, and rely on them to render the service for us. To navigate the world, we necessarily put our faith in others to provide not just physical services but informational and evaluative services on our behalf across a vast array of topics beyond the scope of either our consideration, concentration, or comprehension. Trust, therefore, is the watchword of belief in a modern society.

[8] Adams, T. (2021, May 16). Daniel Kahneman: 'Clearly AI is going to win. How people are going to adjust is a fascinating problem'. The Guardian. https://www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow

This has both potentially encouraging and troubling implications. Troubling because research, detailed in the following section, shows that Americans are increasingly polarised in their political views, and that this extends to perceptions about the credibility of evidence and sources thereof (Lupia, 2013; Druckman & McGrath, 2019). Encouraging because, while recent decades have seen a steep decline in public trust toward once-venerated institutions (such as, for example, the mainstream media; Brennan, 2022), there are still sources, such as NASA, that retain broad credibility across parties (Myers et al, 2013). Using this knowledge, recent research has found that telling people about the aforementioned NASA-endorsed[9] research on scientific consensus increases belief in AGW and support for policy action (van der Linden et al, 2015). Yet, others find that such updating behaviour is asymmetric, such that people are more receptive to information that fits with their existing views of climate change, and this in turn can further exacerbate polarisation on the issue (Sunstein et al, 2016). It is in this context of somewhat conflicting signals that the present research seeks to examine not just whether but also how people update their beliefs in line with new information, and consider the implications of these findings for the likelihood of finding common ground on climate policy.

[9] Link to NASA's *Do scientists agree on climate change?*

3.2 Literature Review

3.2.1 Motivated Thought

A quickly growing literature now exists to study how informational supplements can be used to address individuals' political misperceptions. Grounded in social psychology, much of this research is focused on how motivation affects people's evaluation of new information. In a foundational paper, Kunda (1990) elaborates the way in which motivation (i.e., desire to reach a given conclusion) affects our processing of new information and evidence. She distinguishes two goals that can be seen to drive individuals' consideration of information: *directional* goals (i.e., the desire to arrive at a specific conclusion, particularly one in line with previous beliefs and attitudes) and *accuracy* goals (i.e., the desire to arrive at the *correct* conclusion). While there is considerable debate on how these goals are balanced, the author makes a compelling case that our desire for belief consistency will create a natural preference for our directional goals. This means that, all else equal, we will perceive pro-attitudinal information to be more accurate than counter-attitudinal information, and will be more likely to seek it out, review it, retain it, and make use of it. Further, she contends that individuals will prioritise directional over accuracy goals — and draw conclusions corresponding thereto — to the extent that they can 'construct seemingly reasonable justifications' for drawing such conclusions. That is, only when a directional goal places one's belief so far and obviously afield of his accuracy goal will an adjustment be made in favour of the latter.

This is a critical insight into the mechanism of motivated cognition, and has particularly interesting consequences for the maintenance of political beliefs. Given the focus on self-justification, this understanding of motivated reasoning suggests something of a paradox: more informed voters, who are therefore better able to construct believable justifications for their beliefs, might be less likely to respond to

informational supplements. Some recent studies give credence to this notion. In a particularly interesting example, Kahan and co-authors (2012) show that, among individuals motivated to reject evidence of climate change (US republicans), those who score higher in science literacy are less likely than those with lower science literacy to accept evidence on human-caused climate change (but see Persson et al, 2021). The explanation for this bolsters the view that individuals' self-justifications are important, as those with low science literacy seem to update their beliefs due to an inability to justify not doing so. Conversely, high science literacy allows one to criticise the study on a range of factors (e.g., the strength of its empirical design). Similarly, Gaines and co-authors find that better-informed individuals use interpretation to find meaning in facts that supports views they already hold (Gaines et al, 2007). In another study, researchers found that voters who more habitually read the news and have stronger political beliefs and attitudes are more likely than less habitual news readers to ignore counter-attitudinal information (Knobloch-Westerwick & Meng, 2009).

There are two general mechanisms behind this sort of information processing. First, some believe that preference for pro-attitudinal information — otherwise termed confirmation bias — rests on a heuristic known as confirmatory hypothesis testing. In this view, individuals approach a given question with a specific outcome (i.e., a hypothesis) in mind, and seek out information to confirm that position. That this is believed to be a natural decision-making heuristic is critical, as it means individuals need not be motivated by any particularly strong affiliation with the belief to engage in this sort of biased information search. In a famous example, participants who were told that an individual they were to meet was either introverted or extroverted (in both cases they met the same person) then provided evidence from the interaction ex-ante to support the belief they had been given (Snyder & Swann, 1978). In other words, their perceptions of the interaction were dictated by their initial belief-informed expectation.

In a later test of this hypothesis, individuals were asked to play a logic game known as the Wason selection task, in which one needs to find disconfirming evidence to solve a puzzle (Wason, 1968). Given that those motivated by a belief in favour of a particular conclusion will be more likely to seek confirming evidence (and vice versa), researchers hypothesised that those with sufficient motivation to reject a task rule would perform better in the Wason game because of their predisposition toward disconfirming evidence (Dawson, Gilovich & Regan, 2002). Finding evidence for this claim, the researchers lay out a framework for decision-making that is quite in line with Kunda (1990): in evaluating pro-attitudinal information, individuals will subject that information to an evidentiary standard of *can I accept this information* — whereas, in evaluating counter-attitudinal information, individuals use the standard of *must I accept this information*. This *can versus must* heuristic is the second mechanism that drives differences in propensity to update one's beliefs based on the content of an informational cue.

Most other scholarship on motivated information processing agrees with this general explanation. For instance, Ditto and Lopez (1992) write that, rather than enabling people to draw whatever conclusions they should so desire, 'motivated scepticism' against preference-inconsistent information — or, conversely, motivated desire to believe preference-consistent information — will essentially affect the amount of information required to draw a given conclusion. Again, this suggests that, as we process information, we will subject it to different evidentiary standards based on its valence. In a static information environment, this of course makes sense: if I already have a belief, I will need no additional information in order to hold that belief (self-evident), but at least some information to change it. In a dynamic environment, however, this preference for consistency leads to sub-optimal (e.g., empirically incorrect) beliefs, and these beliefs form the basis for consequential preferences.

Adopting this understanding, the question then becomes how great the difference between the two evidentiary standards to which individuals will subject pro- and counter-attitudinal information is.

3.2.2 Motivated Political Judgment and Belief

The past two decades has seen a surge in experimental and observational research on how voters incorporate information into their political beliefs. Yet, perhaps unsurprisingly, there remains disagreement about the capacity to foster belief change. Bringing the previously mentioned research from social psychology to the study of political behaviour, political psychologists developed a model based on Kunda (1990), which assumes "all reasoning is motivated" (Lodge & Taber, 2000, p.186). Testing their hypotheses, they find strong evidence that individuals subject pro- and counter-attitudinal political information to different standards of scrutiny — and also that these information searches are conducted in line with the model of confirmation bias detailed above (i.e., preferring confirming to disconfirming information) (Taber & Lodge, 2006). Citing Evans and Over (1996), the authors clearly and compellingly explain how this systematic effect of previous beliefs and attitudes places political decision-makers outside the rational Bayesian model of belief updating, which requires that priors are kept independent of the processing of new information (Taber & Lodge, 2006). In the rational model, though priors are a crucial component of judgments, new information is meant to be processed and then used to amend those priors and bring beliefs in line with the best available information. Conversely, the motivated model demonstrates how priors alter the intake, processing, and retention of new information.

There are multiple suggested means through which previous beliefs are thought to affect subsequent processing and evaluation. Often, these models entail two modes of thought, which complement, though can be in contrast to, one another. Such conceptions of judgment are longstanding: "[f]rom the writings of the earliest

philosophers to the present, there has been an almost unbroken belief that human behavior is best understood as the product of two interacting and often competing processes" (Loewenstein, O'Donoghue & Bhatia, 2015, p.55). These *two minds* have been conceptualised variously as containing cognitive and affective roots. For the former, scholars have noted the presence of controlled versus automatic thinking (Shiffrin & Schneider, 1977), famously stylised as *System 1* and *System 2*, where the former pertains to unconscious processes, such as heuristics, that underlie many judgments, and the latter, higher-order deliberations that operate at the level of conscious thought (Stanovich & West, 2000; Kahneman & Frederick, 2002; Kahneman, 2011).

These automatic processes rely heavily on shorthand categorisations to make judgments, and though they are frequently quite adaptive (Gigerenzer, 2008) — allowing for simplification of the vast amounts of information people encounter in navigating daily life — they can also result in the application of poorly adapted heuristics or prejudices, and thus predictable decision errors (see, for example, Tversky & Kahneman, 1974). Moreover, these automatic processes are likely to form the basis of judgments, with people only activating deliberative thought when they fail to form a coherent interpretation (Lieberman, Schreiber & Oschner, 2003). Whereas some view automatic processes as the frequent source of maladaptive updating behaviours, others argue that, in fact, it is with deliberation that individuals are better able to construct self-justifications for arriving at their preferred conclusions (Kahan et al, 2012)[10].

Expanding on the dual-mind paradigm, others contend that beyond simple cognitive shortcuts, automatic processes yield judgments that are more affectively-driven, *hot* with emotion and charged by previous affective judgments that are

[10] But see Bago, Rand & Pennycook (2020).

unconsciously evoked by a given subject (Metcalfe and Mischel, 2000; Lodge & Taber, 2005). In this view, all sociopolitical concepts with which one has ever been familiar, when evoked, are "charged with positive or negative affect" (Burdein, Taber & Lodge, 2006, p.365). Summarising the how these affective cues mediate the effect of information on political judgment, Arcuri and co-authors write that "if the emotions elicited by the new information are coherent with previous evaluations stored in memory and automatically activated, the new information is acquired, accepted, and stored. In contrast, new information that contradicts current spontaneous evaluation is denied, challenged, or simply ignored" (Arcuri et al, 2008, p.372). Moreover, the impact of affective motivation on behaviour is predicted by the intensity of that affect for the object under consideration (Loewenstein, O'Donoghue & Bhatia, 2015).

Whether driven by simple unconscious cognition, affect, or indeed conscious deliberation, the sum of this research demonstrates the frequently occurring, if less consistently predictable, ways in which political judgments are often influenced by directional rather than accuracy motivations. Such effects have been found in voters' evaluations of economic performance, which, though linked to real-world conditions (De Vries, Hobolt & Tilley, 2018), are subject to partisan predispositions to assess performance as better or worse, or form optimistic or pessimistic expectations, depending on whether one's party is in power (Duch, Palmer & Anderson, 2000; Duch & Stevenson, 2011). Or, when evaluations are able to surpass partisan filters, these predispositions are reflected instead in attributions of responsibility and individual sense-making (Tilley & Hobolt, 2011; Bisgaard, 2015). They have been found in policy evaluations, such that one's favourability to a given proposal differs based on whether it has been made by a co-partisan or an out-partisan (Cohen, 2003), or in the use of party cues over evidence-based considerations of the policy itself (Agadjanian, 2021). And they have been found to yield selective learning on political topics, such that people

have more knowledge of relevant facts that confirm their views, and less of those that disagree with them (Jerit & Barabas, 2012).

As in the above paradigm, these biased beliefs are thought to occur through the unequal review of pro- and counter-attitudinal information, and a general resistance to changing beliefs once held (Stanley et al, 2020). Here the distinction between beliefs and attitudes is worth parsing more closely. From a psychological standpoint, an attitude can be thought of as one's broader *feeling* about a given subject, entailing 'some degree of favour or disfavour' (Eagly & Chaiken, 1993), while a belief is used to internalise and order information about a given topic, and need not carry any valence judgment (Fishbein & Ajzen, 1977). For example, one may hold the attitude that climate change is *bad* and should be combatted, while holding the specific belief that, say, a given policy proposal would be ineffectual. From the *unmoved mover* perspective, partisanship itself is a political attitude, driving beliefs to conform thereto. With respect to climate change, as seen in 3.1.1, broader partisan and issue attitudes seem to drive the way in which evidence is considered and ultimately the beliefs that are formed, contributing to a broad divergence in what is seen as credible evidence of AGW. Numerous other examples can be found in the literature on motivated political reasoning, but the takeaway is that people subject political information to different standards of evaluation in line with their predispositions. In turn, these judgments allow for the maintenance of broader attitudes and ideologies, which filter through to the behavioural components of political action and voting.

3.2.3 When Beliefs Do (and Do Not) Change

The previous sections have detailed the extant evidence of motivated political reasoning, and its grounding in psychological models of judgment and decision-making. They impart the axiom that people's existing beliefs affect their subsequent beliefs, through various processes of information-seeking and avoidance, automatic and

affective heuristics, or deliberative self-deception and justification — and that, all else equal, we will be more likely to arrive at conclusions that align with these predispositions. What about when all else is not equal?

Acknowledging the nature of these findings, and often taking as a given people's preference for belief maintenance, a host of modern literature now seeks to evaluate the conditions that *can* lead people to change their beliefs and attitudes in response to new information. Or, to continue with the previous paradigm, to engender a preference for accuracy goals, and answer in the affirmative to the *must I* heuristic. This research has led to sometimes conflicting findings about the factors that most predict belief change. For example, while most agree that elite partisan cues are influential in how partisans view political issues (Levendusky, 2010), some research finds that these cues predominate other factors in the formation of policy beliefs (Agadjanian, 2021). This includes the provision of factual information, the efficacy of which is significantly lower in polarised contexts, in turn leading some to more confidently hold erroneous beliefs (Druckman, Peterson & Slothuus, 2013). In the context of climate beliefs, while consensus messaging can increase belief in AGW (van der Linden et al, 2015), politicisation of the issue along partisan lines significantly decreases the effectiveness of such messages (Bolsen & Druckman, 2018). Indeed, simply making one's partisanship more salient leads to lower reported levels of belief in AGW, and less support for policy to mitigate damages (particularly, but not only, amongst right-leaning individuals) (Unsworth & Fielding, 2014). Looking at this effect another way, Benegal and Scruggs (2018) find that acknowledgements of the human causes of climate change by Republican leaders led to greater acceptance thereof by Republican respondents.

Conversely, others suggest that, while powerful, people can overcome partisan cues and adapt their beliefs in line with new information. This is particularly the case

when partisan messaging fails to resonate with the frame through which people think it is most appropriate to consider a particular issue (Slothuus, 2010) — or when issues are deemed more important and thus worthy of individual fact-finding (Ciuk & Yost, 2016). Moreover, when an in-partisan's policy position is overtly contradicted by newly presented evidence, the persuasive value such information can overcome even a strong partisan cue (Tappin, Berinsky & Rand, 2021). Ultimately this may depend on the extent to which people feel a need to evaluate an issue for themselves (or referring to the previous section, when automatic cues fail to generate a satisfactory resolution and people feel the need to activate a more deliberative process) (Druckman & Nelson, 2003).

Related, in a way, to the effect of partisan cues is the effect that perceived source credibility has on people's propensity to update their beliefs. Related because people's perceptions of credibility are based in part on their own attitudes and identity as well as that of the speaker (Kahan, Jenkins-Smith & Braman, 2011; Metzger & Hartsell, 2020), distinct because there are still information sources, such as NASA, that retain broad credibility across parties (Myers et al, 2013). A great deal of public opinion research corroborates the effect of people's perceptions of source credibility on how they treat information therefrom, intuitively showing that greater credibility makes information more persuasive (Pornpitakpan, 2004). These effects are crucial in policy debates pertaining to scientific evidence that require the translation of expert opinion, particularly in contexts of greater polarisation (Lupia, 2013).

In a similar fashion to how partisan cues mediate the effect of information, where partisan dogmas and individuals' views of specific issues diverge, these views similarly act as a filter for subsequent evaluations. Some research suggests that, if these forces are in opposition, issue-attitudes are a more powerful predictor of ultimate evaluations than partisanship (Mullinix, 2016). One explanation for this may be that,

because of the ways in which people outsource beliefs on low-importance issues, divergence marks cases people have felt the *need to evaluate* on their own. Ultimately the relationship between issue-attitudes and partisanship is a two-way street, and high levels of divergence there between can help lower the partisan filter (Highton & Kam, 2011).

Though hardly settled, the sum of this research has greatly advanced our understanding of the contextual factors that most influence people's propensity for belief change. As explained in the sections to follow, these lessons form the basis of the information treatments that comprise the present experiment.

3.2.4 Corrected Misperceptions and Their Consequences

The model of motivated political reasoning becomes increasingly troubling when one reconciles it with the previously mentioned issue of widespread misperceptions. It suggests that not only are individuals reluctant to incorporate new information and evidence into their beliefs, but simultaneously that the beliefs being defended are often uninformed, or worse, based on mis- or dis-information. Thus, much recent research on belief change focuses specifically on the correction of misperceptions, and how they affect subsequent beliefs.

Two particularly prominent and contradictory articles on the so-called 'backfire effect' have framed much discussion about how, if at all, informational supplements can lead to an alignment of beliefs and evidence. In the original article, the main argument posited — and then supported with experimental evidence — is that correcting misinformation can lead those who hold contrary beliefs to not just not change their beliefs, but in fact to deepen their support of the incorrect belief (Nyhan & Reifler, 2010). In a first example, the authors show that individuals who believed evidence presented by the Bush Administration of weapons of mass destruction (WMD) in Iraq not only did not update their beliefs based on a correction, but strengthened their prior, incorrect belief. These corrections were conducted in an

experimental context in which participants read fabricated news reports containing incorrect information, and in which the treatment group would then read a correction of that misinformation. The difference between groups (i.e., those who did and did not receive the correction) illustrates the effect of factual corrections — which the authors show not only do not lead to belief convergence, but can lead to belief retrenchment.

Yet, in an authoritative response to this work, Wood and Porter (2019) conduct four studies across thirty "commonly misconstrued" (p.138) policy areas to look for occurrences of this effect. In 29 of the 30 evaluated domains, the authors find no evidence of backfire, but rather find participants "capable of heeding the correction and bringing their beliefs in alignment with the facts" (p.139). In the single instance where backfire did occur (in a replication of the WMD prompt from the original) the authors note that simplifying the survey item caused the effect to disappear. In other words, at least part of the results found in the original paper could be attributable to the operationalisation, rather than an underlying mechanism. Nonetheless, though corrections may not backfire, the authors did observe important differences in belief and attitude change: while all respondents' beliefs moved in the direction of alignment with the facts presented, partisanship and individuals' ideology were strong mediators of the extent of that change. This aligns with other research on how the effectiveness of evidence on climate consensus is stunted by political predispositions (Bolsen & Druckman, 2018).

Beyond examinations of backfire, a number of other recent studies focus on the prevalence and content of public misperceptions, and the role of motivated information processing in the construction and reporting of beliefs. Using a similar design, Weeks and Garrett (2014) examine how receptive individuals are to political rumours about favoured and opposing candidates. They find, in line with the model of motivated reasoning outlined above, voters are more likely to believe negative rumours about

opposing candidates than favoured ones. However, the authors also find that correcting the misinformation of the rumours is effective for beliefs about both pro- and counter-attitudinal rumours, and there is no evidence of backfire. With respect to broader attitudes, however, the effect of corrected information is rather different than beliefs. Employing a belief perseverance model (Anderson, Lepper & Ross, 1980), Bullock (2006) finds that discredited misinformation still has lingering attitudinal effects, such that the original (incorrect) claim still affects participants' attitudes even after correction. This is in line with research on so-called 'belief echoes', or the continued-influence effect (CIE), which argue that the attitudinal effects of misinformation linger well beyond the point of discrediting (Johnson & Seifert, 1994; Thorson, 2016; Ecker & Ang, 2019).

Drawing conclusions from this body of work poses challenges, as there is, on either hand, compelling evidence for and against evidence-based belief updating. Ironically, this may engender beliefs about evidence-based beliefs that are themselves counter to the best evidence. One clear lesson bridging these competing findings is that context matters, and strong partisan and issue priors mediate the effects of high-quality information. Moreover, when beliefs do align with information, people's interpretations thereof can still lead to divergence in how they make sense of that information.

3.2.5 Research Puzzle

Partial or conflicting answers to the question of how people adapt incorrect beliefs to new facts demonstrate the need for further study. One of the limitations of some past research is the difficulty of separating facts from opinions, such that people may not feel a correction is firmly grounded in provable evidence (Walter & Salovich, 2021). Consider, for instance, some of the cases used in the previous studies on backfire effects. Figure 3.1 below shows eight examples of misconceptions (column 1) the authors use to test for the effectiveness of corrections and potential backfire.

Figure 3.1: Classifying corrections in Wood & Porter (Study 1)

Incorrect statement	Speaker	Correction given	Outcome statement	Agreement without correction (1 to 5)	Reduction in agreement with correction	Reduction in agreement (%)
"Between 88 and 92 people a day are killed by guns in America. It's the leading cause of death for young black men, the second leading cause for young Hispanic men, the fourth leading cause for young white men. This epidemic of gun violence knows no boundaries, knows no limits, of any kind." (2015)	Hillary Clinton	In fact, according to the FBI, the number of gun homicides has fallen since the mid 1990s, declining by about 50% between 1994 and 2013.	The number of gun homicides is currently at an all time high	3.93	-1.25	-31.8%
"In 1980, there were 500,000 people behind bars in America. Today there are 2.2 million. Over the last few decades, we've also locked up more and more nonviolent drug offenders than ever before, for longer than ever before. And that is the real reason our prison population is so high." (2015)	Barack Obama	In fact, according to the Federal Bureau of Justice Statistics, only 20% of prisoners in the US committed a nonviolent drug offense. Over half of all prisoners were convicted of violent crimes.	Drug convictions are the main cause of the high prison population	4.49	-0.99	-22.0%
"We need to get back into the habit of actually rewarding workers with increases in their paychecks... Warren Buffett has said it, but so have a lot of other people. There's something wrong when hedge fund managers make more, and pay less in taxes, than nurses or truck drivers." (2015)	Hillary Clinton	In fact, according to the Bureau of Labor Statistics, the average hedge fund manager pays about 20 times as much income tax as the average truck driver or nurse.	Nurses and truck drivers pay more in taxes than hedge fund managers	3.99	-0.78	-19.5%
"Today, women make up about half our workforce. But they still make 77 cents for every dollar a man earns. That is wrong, and in 2014, it's an embarrassment. A woman deserves equal pay for equal work... It's time to do away with workplace policies that belong in a "Mad Men" episode." (2014)	Barack Obama	In fact, a report for the Department of Labor found that the Wage Gap claim does not take into account differences in occupation, position, education, job tenure or hours worked per week. After adjusting for these factors, women earn 95 cents for every dollar men earn.	Women are paid less than men for the same work, mainly because of discrimination	4.53	-0.64	-14.1%
"Defense spending has fallen dramatically on President Obama's watch, and our forces have been reduced and suffered readiness challenges even as threats abroad increase. Our military therefore needs a serious program of reinvestment and modernization." (2015)	Marco Rubio	In fact, according to the Defense Department, defense spending under President Obama is higher than it was under President George W. Bush.	Defense spending under President Barack Obama is lower than it was under President Bush	2.76	-1.06	-38.4%
"When Mexico sends its people, they're not sending their best. They're sending people that have lots of problems, and they're bringing those problems to us. They're bringing drugs. They're bringing crime. They're rapists." (2015)	Donald Trump	In fact, according to the Congressional Research Service, undocumented/illegal immigrants commit crimes at a lower rate than the general population.	Undocumented / illegal immigrants commit crimes at a higher rate than most Americans	1.27	-0.21	-16.5%
"The number of law enforcement officers killed as a result of violence has been on a precipitous upswing. If the police are intimidated, if they are scared, if they are not willing to do their jobs, we know the result. The result is the loss of life. The result is rising crime." (2015)	Ted Cruz	In fact, according to public records, homicides of law enforcement officers have been declining for decades. Fewer police officers were killed in 2015 than any year since the 1890s (sic?).	The number of police officers killed in the line of duty is rising fast	2.08	-0.67	-32.2%
"Our taxes–we just put in a plan the other day–we're going to reduce taxes tremendously because we have the highest tax rate anywhere in the world and our middle class is being absolutely destroyed." (2015)	Donald Trump	In fact, according to the Organization for Economic Cooperation and Development, taxes account for only about 25% of the US economy–the second smallest share among all advanced nations.	U.S. tax rates are the highest in the world	1.81	-0.19	-10.5%

Source: Wood & Porter (2019, Appendices A.1 & A.2); Author's calculations

Alongside these misperceptions are the relevant corrections issued (column 3) and the outcome statement used to measure agreement (column 4). The design is between-subjects, such that some participants indicate their agreement (Likert scale from 1 to 5) with the incorrect statement without receiving a correction (column 5). In the treatment, other participants read the same statements before being issued the correction and then indicating their own belief with the original statement. The reduction in agreement from no correction to correction is shown in column 6, with the percentage change in column 7.

As is often the case with political communications, the original statements are a blend of factual information and interpretation. Drawing from these, consider just a few differences between the misconceptions presented, the corrections, and the outcome statements used to measure agreement. For example, the case with the greatest reduction in agreement, a statement by Marco Rubio:

Original statement: "Defense spending has fallen dramatically on President Obama's watch, and our forces have been reduced and suffered readiness challenges even as threats abroad increase. Our military therefore needs a serious program of reinvestment and modernization." (2015)

Correction given: In fact, according to the Defense Department, defense spending under President Obama is higher than it was under President George W. Bush.

Outcome statement: Defense spending under President Barack Obama is lower than it was under President Bush (Wood & Porter, 2020, Appendix A.1)

There is little to take issue with in this example. The statement makes a clear assertion that defence spending had fallen under President Obama, with the necessary implication that it was higher under his predecessor, President Bush. The correction states clearly that defence spending was in fact higher under Obama than under Bush, and the outcome directly measures belief regarding whether defence spending had fallen under Obama. The corresponding reduction in belief with the correction is -38.4%. Conversely, consider an example perhaps less-grounded factual debate:

Original statement: "Today, women make up about half our workforce. But they still make 77 cents for every dollar a man earns. That is wrong, and in 2014, it's an embarrassment. A woman deserves equal pay for equal work... It's time to do away with workplace policies that belong in a "Mad Men" episode." (2014)

Correction given: In fact, a report for the Department of Labor found that the Wage Gap claim does not take into account differences in occupation, position, education, job tenure or hours worked per week. After adjusting for these factors, women earn 95 cents for every dollar men earn.

Outcome statement: Women are paid less than men for the same work, mainly because of discrimination (Wood & Porter, 2020, Appendix A.1)

Parsing these quotes, there are a few potential issues with this example. First, the original quote does not explicitly say the difference in pay is solely on account of discrimination (even if the statement *equal pay for equal work* may imply so). Second, there are contradictory analyses on the gender wage gap and the relative contribution of discrimination thereto[11], and many of these differences stem from conflicting ways of conceptualising discrimination itself (e.g., strictly lower pay for the same work versus devaluation of women's labour). Finally, the correction still notes a $0.05 gender gap in pay, whereas we are left to suppose about the relative contributions of the other factors listed. This likely bears on agreement with the outcome, which measures belief in whether discrimination is the greatest contributor to the gender wage gap. The corresponding reduction in belief with correction here is -14.1%.

These examples are in no way meant to call into question the contribution of the original study, nor to imply that the corresponding differences in belief changes are a result of this variance between the topics used. Rather, it is simply meant to illustrate the frequent comingling of facts and interpretation in past research on corrective information. Of course, part of the reason for research designs of this nature is to reflect the reality of political communication: rarely are conversations about politics and

[11]At least in the view of the former chief economist for the US Department of Labor. (Jones, J. (2021, March 19). *5 Facts About the State of the Gender Pay Gap | U.S. Department of Labor Blog*. Blog.dol.gov. https://blog.dol.gov/2021/03/19/5-facts-about-the-state-of-the-gender-pay-gap)

policy strictly fact-based, but rather a blend of opinion and evidence. Nonetheless, it seems likely that, as a result of such designs, past research may fail to address directly the question of factual receptivity, *per se.*

One key difficulty that may arise from mixing facts and opinions is that people themselves often have trouble differentiating between the two (Merpert et al, 2018). For example, a study run by the Pew Research Center found that only 26% of people correctly classify the whole of a set of five factual statements (e.g., spending on Social Security, Medicare, and Medicaid make up the largest portion of the US federal budget) versus five opinion-based statements (e.g., government is almost always wasteful and inefficient) (Pew, 2018). In turn, corrections that incorporate both fact- and opinion-based elements are less effective on the whole than strict fact checks (Walter & Salovich, 2021). One explanation for this is that people deem the corrections subjective or unprovable.

The present research, then, is focused on updating behaviour in response to a more evidently provable (indeed, a physical) fact. In the experiment that follows, I exploit a widespread misperception about whether NG and MG are different substances (they are not), and provide information about the physical properties of these gases. As mentioned before, this context is particularly well-suited to studying updating behaviour not only because it focuses on a demonstrable fact, but also because the original gap in beliefs between NG and MG allows me to observe not just whether, but also how, people update their beliefs — and how the direction of updating corresponds to broader attitudes.

The primary research questions for the experimental analysis that follows are:

RQ1: Are misperceptions regarding NG and MG in line with past research?

RQ2: How responsive are these misperceptions to a neutrally-framed and credibly-sourced factual correction?

RQ3: How (if at all) do source credibility, partisan cues, and strong climate framing alter the efficacy of the informational correction?

RQ4: How do broader partisan and climate attitudes affect updating behaviour?

Each of the hypotheses that will be evaluated draw from the research reviewed above. This includes findings that show a (perhaps modest) willingness to update incorrect beliefs, particularly when the evidence presented is well-sourced, neutrally-framed, and factual. Thus I look not only at whether and how beliefs update, but also how this is affected by the presentation of evidence in the correction. The hypotheses are as follows:

H1: Information cues lead to convergence of NG and MG belief

H2: Changes to the correction yield less updating, with partisan cues and climate framing lowest in belief convergence

H3: Partisan differences in the direction of NG and MG belief changes in line with broader attitudes

H4: Climate attitudes moderate the effect of information on convergence in the climate-framed treatment

As I explain in the experimental design in 3.3.1, the first treatment focuses on whether and how individuals update their NG and MG beliefs in response to a the neutral, sourced fact check, while each of the subsequent conditions tests the sensitivity of updating behaviour to one of the facets of the treatment (that is, either the sourcing or framing of the evidence). The second hypothesis, then, posits that the neutral presentation of evidence will be the most effective at yielding belief convergence. In converging their beliefs, the expectation is that — given the differences in broader climate and energy attitudes detailed above — Democrats will be more likely to downwardly update their NG beliefs, while Republicans would be more likely to increase their favourability toward MG. Finally, the invocation of partisan and climate cues are expected to neutralise the effectiveness of the treatment.

3.3 Experimental Design & Methods

3.3.1 Study Design

The multi-armed experiment that follows aims to assess the responsiveness of a misperception to high-quality information. Specifically, I am interested in both the extent of convergence between NG and MG beliefs, as well as the direction of belief change for each (i.e., updating favourability toward NG downward or MG upward). I consider propensity for belief change given four slightly modified corrections, each reflecting factors shown in past research to alter the efficacy of new information. These information treatments are as follows:

Sourced Fact Information Treatment (T1): The first treatment, T1, serves as a framing-neutral condition against which the subsequent treatments are compared. Participants first indicate their favourability toward various energy sources (0-10), including NG and MG. After answering, they then receive the informational treatment, which reads:

Many Americans hold different views toward methane gas and natural gas. However, in composition these gases are essentially the same, as natural gases are comprised of between 70-95% methane. As the US Department of Energy describes it, natural gas "is primarily methane with small amounts of other hydrocarbons."

According to NASA, methane gas and natural gas are the same in terms sustainability and efficiency, with both posing the same health and environmental risks. Thus, though the labels are different, the gases themselves are the same.

Following this treatment, participants proceed to the following page, which reads:

Previously, your favorability[12] toward methane gas was [link to previous answer] and your favorability toward natural gas was [link to previous answer]. Based on the knowledge that natural gases are methane gases, and both energy sources are the same in terms of sustainability and efficiency, how would you rate your favorability toward each these energy sources? (0 is very unfavorable and 10 is very favorable)

After indicating their updated (or not updated) favourability, participants then answer a short series of political and demographic questions and the study concludes.

Unsourced Fact Information Treatment (T2):
This treatment considers how perceived source credibility may affect updating behaviour. It is identical to T1, except that references to both the DOE and NASA are removed. The modified text then reads:

Many Americans hold different views toward methane gas and natural gas. However, in composition these gases are essentially the same, as natural gases are comprised of between 70-95% methane. Natural gas "is primarily methane with small amounts of other hydrocarbons."

Methane gas and natural gas are the same in terms sustainability and efficiency, with both posing the same health and environmental risks. Thus, though the labels are different, the gases themselves are the same.

[12] Note that, as the study was run in the US, spellings within the experiment are consistent with US English conventions.

Information + Partisan Cue Treatment (T3):
This treatment considers how partisan cues affect updating behaviour. Participants follow the same experimental pathway as in T1, with the following addition made to the information treatment:

In recent years, natural gas production has become an issue of political disagreement. Some Democratic leaders have moved to cut usage of natural gas as a source for electricity. Conversely, some Republican leaders have taken a firm stance in favor of increasing Americans' use of natural gas, which they termed "Freedom Gas" in a 2019 press release.

Information + Climate Cue Treatment (T4):
This treatment considers how evoking broader climate attitudes affects updating behaviour. As in T2, participants receive the same information treatment from T1, with the below addition, which seeks to frame the issue more starkly along lines of climate change beliefs:

In recent years, natural and methane gas production has been scrutinized for its negative impacts on climate change. According to NASA Administrator Bill Nelson, "reining in methane emissions is key to limiting global warming." A 2018 paper published in the journal Science found that methane leaks during the production process have doubled the climate impact of natural gas "causing warming on par with carbon dioxide (CO2)-emitting coal plants for 2 decades." Some critics argue that limiting methane gas production will be key to meeting the climate goals set out in the 2016 Paris Climate Accords.

For the experiment, I recruit participants (n = 633) from the online panel provider Lucid. After providing informed consent, participants are randomised into one of the four experimental treatments detailed above. Lucid's samples balance participant attributes against benchmarks from US census data along dimensions of age, gender, and political partisanship. Past research on data quality in online responses has sought to replicate a range of canonical past studies, finding that Lucid samples yield treatment effects in line with the originals and are suitable for social science experimentation (Coppock & McLellan, 2019).

3.3.2 Methods

Two main outcome variables are used to evaluate these hypotheses: the extent of belief convergence and the direction of underlying belief changes. In order to account for initial belief gaps, the measure of convergence is normalised such that it calculates the difference between the original gap and the final gap, divided by the original gap. This

allows for comparisons of the extent of individuals' relative belief change (i.e., what proportion of the gap is bridged by the treatment) based on experimental condition.[13]

Let $\hat{\mu}_{Nt}$ be the mean[14] of an individual's beliefs about NG and $\hat{\mu}_{Mt}$ their beliefs about MG at timepoint *t,* or prior to treatment. Let $\hat{\mu}_{Nu}$ and $\hat{\mu}_{Mu}$ be mean beliefs at timepoint *u*, or after treatment. Beliefs on NG and MG are thus recorded to have converged if and only if: $|\hat{\mu}_{Nt} - \hat{\mu}_{Mt}| > |\hat{\mu}_{Nu} - \hat{\mu}_{Mu}|$ whereas beliefs diverge if and only if the opposite holds: $|\hat{\mu}_{Nt} - \hat{\mu}_{Mt}| < |\hat{\mu}_{Nu} - \hat{\mu}_{Mu}|$. To consider a simple example, assume the individual's NG and MG beliefs at time *t* are 6 and 2, respectively. These beliefs at time *u* are 4 and 3 respectively. This individual has clearly updated their beliefs in the direction of information, or in other words has converged their beliefs, such that $|6 - 2| > |4 - 3|$. The absolute convergence for this individual is the difference between their pre-treatment gap and post-treatment gap, calculated as $|\hat{\mu}_{Nt} - \hat{\mu}_{Mt}| - |\hat{\mu}_{Nu} - \hat{\mu}_{Mu}| = 3$. Relative convergence is calculated as $\frac{|\hat{\mu}_{Nt} - \hat{\mu}_{Mt}| - |\hat{\mu}_{Nu} - \hat{\mu}_{Mu}|}{|\hat{\mu}_{Nt} - \hat{\mu}_{Mt}|}$. Thus, this individual's relative convergence is $\frac{|4-1|}{|4|} = 0.75$, or 75% belief convergence.

Directional changes are considered within the contexts of both specific gas beliefs. That is, NG beliefs are compared to NG beliefs, and MG to MG, between timepoints. Consider a pair of individuals, one Democrat, D, and one Republican, R. Assume these individuals hold the same pre-treatment beliefs, that is $|\hat{\mu}_{DNt} - \hat{\mu}_{RNt}| - |\hat{\mu}_{DMt} - \hat{\mu}_{RMt}| = 0$ and both fully converge their NG and MG beliefs, such that

[13] The reason to focus on relative rather than absolute belief convergence is best illustrated with an example. Consider two individuals: A, whose original gap is 8, and B, whose original gap is 3. Assume that, following treatment, A ends with a belief gap of 4 points, B with a gap of 1. An absolute measure would view the former case as yielding greater updating, as A's beliefs converged by 4 points, compared to B's 2. Yet A's gap has only reduced by 50%, whereas B's has reduced by 67%. Moreover, A ends with greater disparity between NG and MG beliefs than B began with. Clearly, then, absolute convergence is not the best measure for capturing convergence of beliefs.

[14] I use $\hat{\mu}$ here under the assumption that people's favourability ratings are essentially an estimate of their *true beliefs* at a given time, but that these estimates may be somewhat variant to time or other conditions.

$|\hat{\mu}_{DNu} - \hat{\mu}_{DMu}| = 0$ and $|\hat{\mu}_{RNu} - \hat{\mu}_{RMu}| = 0$.[15] Directional polarisation on a specific gas, for example NG, will have occurred if and only if: $||\hat{\mu}_{DNu} - \hat{\mu}_{RNu}| - |\hat{\mu}_{DNt} - \hat{\mu}_{RNt}|| > 0$ — or more simply if subtracting the pre-treatment gap from the post-treatment gap gives a positive value. For example, assume $\hat{\mu}_{DNt} = 6$ and $\hat{\mu}_{DNu} = 3$, while $\hat{\mu}_{RNt} = 6$ and $\hat{\mu}_{RNu} = 5$. The result of the expression is then $|3 - 5| - |6 - 6| = 2$. In this case, the post-treatment gap in D and R's NG beliefs is greater than the pre-treatment gap; thus, their NG beliefs have polarised in response to the information treatment. The same logic holds for the examination of MG beliefs. Because the data are individual, and thus cases can only contain observations for *either* R or D beliefs, polarisation is evaluated by considering whether an individual respondent updates their NG beliefs upward or their MG beliefs downward, and then comparing these propensities between parties. For example, an individual is classified as having updated their NG beliefs downward if and only if $\hat{\mu}_{Nt} > \hat{\mu}_{Nu}$, and so forward.

As mentioned previously, the main focus within this outcome is whether NG beliefs adjust down or MG beliefs adjust up (i.e., $\hat{\mu}_{Mu} > \hat{\mu}_{Mt}$), and how partisan attitudes affect this direction. Recalling the belief-attitude framework sketched in 3.2.2, the idea here is that — though people's specific beliefs about NG and MG may be open to changing based on the evidence presented — broader party and issue-attitudes will predict the direction in which they adjust their beliefs. While the measure of beliefs employed is indirect (that is, rather than asking people what percent of NG they believe is comprised of methane, for example, participants instead report favourability toward the two gases), convergence or lack thereof serves as the means of evaluating updating, while the predictive capacity of attitudes in belief change is reflected by how those beliefs update.

[15] Thus, for the previous measure, both individuals exhibit 100% belief convergence.

3.4 Results

3.4.1 Existing Gap in Beliefs

To begin, I first check the extent to which participants hold initial misperceptions about NG and MG, and whether the gap observed in beliefs is in line with past research. As a refresher, past studies have found that, despite some partisan differences in favourability ratings, Americans across the political spectrum hold generally favourable views toward NG, and unfavourable views toward MG. Table 3.1 below presents mean ratings for both gases, as well as the *favourability gap* from NG to MG, split by respondents' preferred political parties. Recall that responses are recorded on 0-10 scales, where lower values indicate less favourable attitudes and vice versa.

Table 3.1: Mean favourability towards NG and MG by party

	All	**Republicans**	**Democrats**	**Independents**
Natural gas	6.63	7.45	6.21	6.33
Methane gas	4.09	4.54	3.86	3.94
Favourability gap	38%	39%	38%	38%
N	633	193	234	206

Note: Favourability is recorded on a 0-10 Likert scale from *very unfavourable* to *very favourable*

In line with past research, I find that respondents are indeed more favourable towards NG ($\mu = 6.63$) than MG ($\mu = 4.09$), with a mean drop in favourability of 38%. Moreover, while there are partisan differences in beliefs — with Republicans generally more favourable toward both labels — participants of all partisan identities hold generally favourable views of NG and generally unfavourable views of MG, with similar perception gaps there between.

I next check whether these misperceptions are consistent across experimental conditions, in order to evaluate the comparability of subsequent measures of updating behaviour across contexts.

Table 3.2: Mean favourability towards NG and MG by condition

	All	**Sourced Fact**	**Unsourced Fact**	**Partisan Cue**	**Climate Cue**
Natural gas	6.63	6.60	6.73	6.59	6.59
Methane gas	4.09	4.19	4.10	3.99	4.10
Favourability gap	38%	36%	39%	39%	38%
N	633	156	161	161	155

Note: Favourability is recorded on a 0-10 Likert scale from *very unfavourable* to *very favourable*

The results in Table 3.2 show a remarkable similarity in responses between conditions. This is critical, as it allows for attribution of any subsequent changes to differing treatment effects, rather than underlying differences in original beliefs. Thus, I conclude the misperceptions observed in past research are present in the current context, and move forward to examine the main research questions.

3.4.2 Do Beliefs Converge?

The first of the primary analyses focuses on whether participants update these misperceptions in line with the informational treatments, and if and how this varies by condition. To evaluate these questions, I use the previously explained measure of belief convergence. As a reminder, this measure shows the percent of one's initial belief gap that is bridged by the factual correction. Thus, these analyses focus only on participants who hold an existing misperception (n =489), excluding participants with no pre-treatment gap in NG and MG beliefs (n = 144).[16] Table 3.3 presents mean belief convergence by condition.

[16] Appendix A.3.2 tests whether the likelihood of holding accurate pre-treatment perceptions (i.e., no belief gap) differs by condition, finding no significant differences there between.

Table 3.3: Mean belief convergence by condition

	All	Sourced Fact	Unsourced Fact	Partisan Cue	Climate Cue
Belief convergence	59.3%	71.0%	61.1%	46.4%	59.8%
N	489	114	125	127	123

The first result that stands out is that, on the whole, participants do align their beliefs on NG and MG in response to information. Across the sample, roughly 59% of the belief gap is corrected. In line one of the expectations set out in H2, the neutral and well-sourced framing of information in the sourced fact (T1) condition yields the greatest updating ($\mu = 71\%$), with less belief convergence across all subsequent conditions. In particular, partisan cues (T3) yield by far the least convergence amongst all conditions ($\mu = 46.4\%$). Conversely, the climate framing of T4 ($\mu = 59.8\%$) does not affect updating behaviour as negatively as T3, but is rather more in line with the source effects of T2 ($\mu = 61.1\%$), which contradicts *H2* expectations.

To test whether these differences reveal a statistically notable difference, I estimate a linear regression model of experimental condition on belief convergence. These results, presented in Table 3.4 below, shows that only the partisan cues of T3 yield significantly less belief convergence than the other conditions.

Table 3.4: Effects of experimental treatments on belief convergence

	Belief convergence
(Intercept = Sourced fact)	0.71 ***
	(0.00)
Condition = Unsourced fact	-0.10
	(0.34)
Condition = Partisan cue	-0.25 *

	(0.02)
Condition = Climate cue	-0.11
	(0.28)
N	489
R2	0.01

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

I next test if the inclusion of key covariates effects these results, in particular whether convergence is influenced by partisanship or climate beliefs. Table 3.5 presents three further models of the effects of treatments on belief convergence, with the addition of partisanship (Model 1), age and gender (Model 2), and belief that "climate change is a serious issue that the government should address" (Model 3). The negative effects of partisan cues are consistent across model specification, while neither political party nor the broader climate attitude toward affect the extent of belief convergence, though respondent age does, such that older respondents converge their beliefs considerably more.

Table 3.5: Effects of experimental conditions on belief convergence with covariates

	Model 1	Model 2	Model 3
(Intercept = Sourced fact)	0.69 ***	0.43 ***	0.22
	(0.00)	(0.00)	(0.16)
Condition = Unsourced fact	-0.10	-0.12	-0.13
	(0.32)	(0.23)	(0.22)
Condition = Partisan cue	-0.25 *	-0.25 *	-0.24 *
	(0.02)	(0.02)	(0.02)
Condition = Climate cue	-0.11	-0.11	-0.11
	(0.27)	(0.29)	(0.28)
Party ID = Independent	0.06	0.09	0.12

	(0.49)	(0.29)	(0.18)
Party ID = Republican Party	-0.00	-0.02	0.05
	(0.99)	(0.85)	(0.57)
Age = 30-50		0.31 **	0.31 ***
		(0.00)	(0.00)
Age = 50-65		0.32 **	0.32 **
		(0.00)	(0.00)
Age = 65+		0.51 ***	0.52 ***
		(0.00)	(0.00)
Female		-0.00	-0.01
		(0.95)	(0.85)
Serious issue			0.02
			(0.06)
N	489	489	489
R2	0.01	0.06	0.06

Note: For partisanship, the reference category is Democratic Party. For age group, the reference category is 18-30. Gender is coded as: 1 = male, 2 = female. Serious issue reports agreement (0-10) with the statement *Climate change is a serious issue that the government should address*. *** p < 0.001; ** p < 0.01; * p < 0.05.

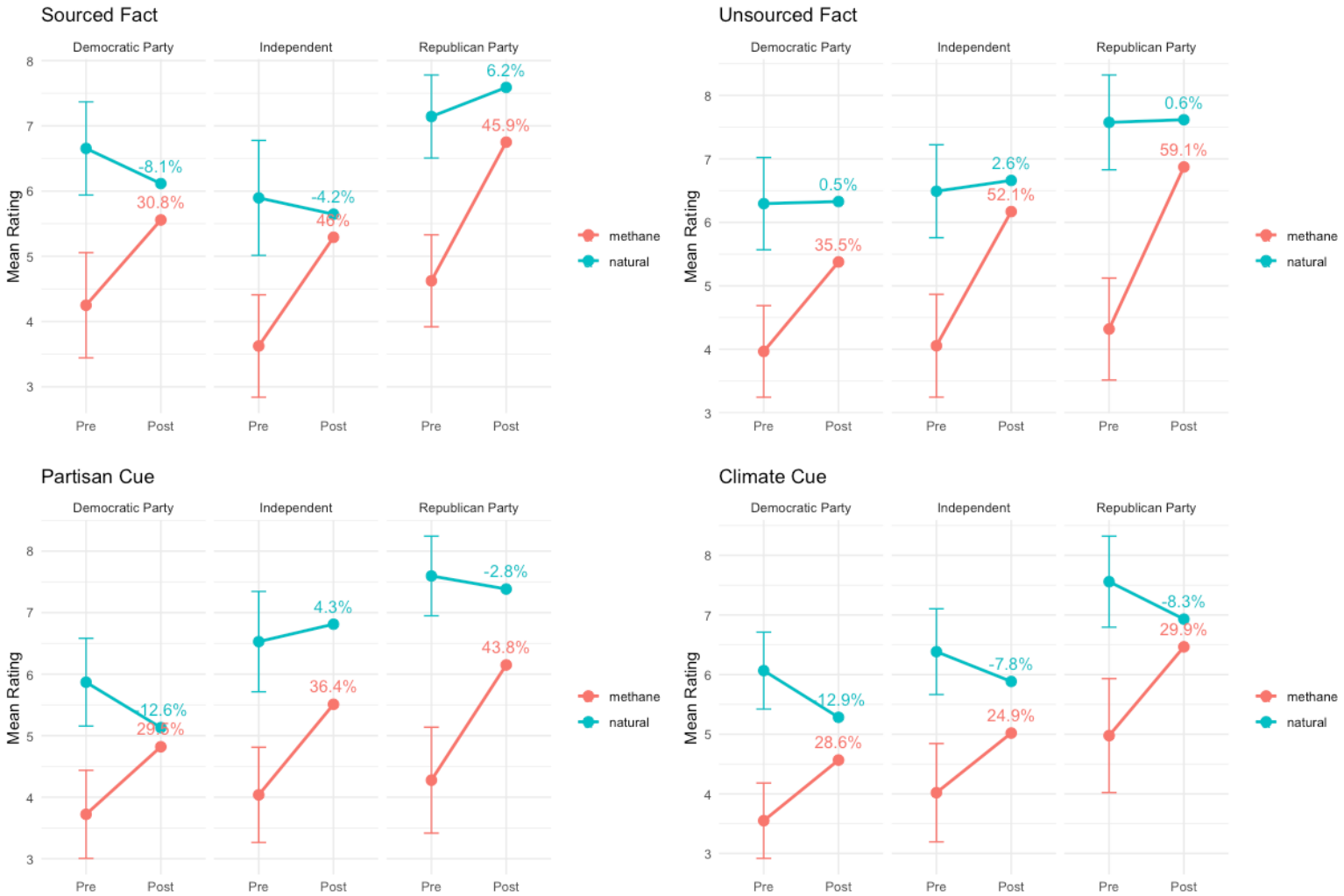
To summarise the results so far, I have found that beliefs on NG and MG do generally converge following correction, in line with the expectations set out in H1. This effect holds regardless of individuals' partisan preferences. Partisan cues in T3 significantly erode belief convergence, however, there is only partial support for H2, as effects of the climate frame in T4 are not of the same magnitude nor significance. Moreover, given that the broader climate attitude pertaining to the seriousness of climate change does not affect convergence, I exclude the (also insignificant) interaction model of climate attitudes and condition from the main text (though it can be found in A.3.3), and reject H4.

3.4.3 Direction of Belief Change

Having observed that beliefs converge in response to information, what is the direction of change for underlying NG and MG beliefs? Based on past research showing partisan differences in fossil fuel attitudes and asymmetries in updating behaviour, H3 hypothesised that partisans would differ in the direction of their belief changes. Or in other words, that, in updating their beliefs, specific gas beliefs would polarise. Figure 3.2 on the following page depicts changes to NG and MG favourability ratings from pre- to post-treatment, split by treatment and political party.

There are a few results that stand out from these plots. First, perhaps unfortunately, participants of all parties tend to update their MG beliefs upward more than they adjust their NG beliefs downward. This may reflect greater familiarity with, and thus more belief rigidity for, the NG label, as *methane gas* is not a frequently used term. However, there are evident differences in directional updating based on both partisanship and experimental treatment. In particular, the downward adjustment of NG beliefs by Democrats (though lower than the upward adjustment of MG beliefs) is significantly greater than either Republicans or Independents. Indeed, in converging their beliefs, Republicans in the sourced (T1) and unsourced fact (T2) conditions, and Independents in the unsourced fact (T2) and partisan cue (T3) conditions, update both MG *and* NG beliefs upwards (though the former change is much greater than the latter). Finally, the climate framing yields the most similar behaviour across parties, with all respondents recording their greatest level of downward NG adjustment out of any of the conditions.

Figure 3.2: Direction of belief changes by treatment and party



These results suggest considerable differences in directional updating based on partisanship and treatment — are these differences statistically meaningful? To test this I estimate a series of logistic regressions of partisanship and experimental condition on direction of change, separately and together and with and without covariates. As explained in 3.3.2, directional changes are evaluated by looking at specific underlying changes to see if NG beliefs adjust downward ($\hat{\mu}_{Nt} > \hat{\mu}_{Nu}$), or MG beliefs adjust upward ($\hat{\mu}_{Mu} > \hat{\mu}_{Mt}$). Thus, the outcome variables are binary, classifying respondents by whether a specific belief moves up or down. Tables 3.6 and 3.7 present the first such tests for partisan differences in downward NG adjustment and upward MG adjustment, respectively.

Table 3.6: Effect of partisanship on downward NG adjustment

	Belief change
(Intercept = Democratic Party)	-0.45 ***
	(0.00)
Party ID = Independent	-0.44 *
	(0.03)
Party ID = Republican Party	-0.74 ***
	(0.00)
N	633
Pseudo R2	0.03

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table 3.7: Effect of partisanship on upward MG adjustment

	Belief change
(Intercept = Democratic Party)	0.02
	(0.90)

Party ID = Independent	0.16
	(0.41)
Party ID = Republican Party	0.59 **
	(0.00)
N	633
Pseudo R2	0.02

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

These results confirm the trends observed above: there are significant differences in directional updating by political party. Or, in other words, there is belief polarisation for the underlying NG and MG beliefs. Specifically, while all respondents are more unlikely than likely to update NG beliefs downward, Democrats are much more likely than Republicans (and somewhat more likely than Independents) to do so. Similarly, while Democrats are slightly more likely than unlikely to update their MG beliefs upward, there are again partisan differences such that Republicans are significantly more likely to increase their favourability toward MG. Both of these findings are in line with expectations set out in H3, and similar to polarisation observed in past research (Sunstein et al, 2016). Following the same procedure, Tables 3.8 and 3.9 consider whether there are differences in updating behaviour based on treatment.

Table 3.8: Effect of condition on downward NG adjustment

	Belief change
(Intercept = Sourced fact)	-1.13 ***
	(0.00)
Condition = Unsourced fact	0.16
	(0.55)
Condition = Partisan cue	0.36
	(0.15)

Condition = Climate cue	0.75 **
	(0.00)
N	633
Pseudo R2	0.02

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table 3.9: Effect of condition on upward MG adjustment

	Belief change
(Intercept = Sourced fact)	0.34 *
	(0.04)
Condition = Unsourced fact	0.18
	(0.42)
Condition = Partisan cue	-0.20
	(0.38)
Condition = Climate cue	-0.35
	(0.13)
N	633
Pseudo R2	0.01

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

In line with what could be seen in Figure 3.1, participants are also significantly more likely to adjust their NG beliefs downward in response to the climate framing in T4. This could be because the additional mentions in that treatment of the global warming consequences of methane-based gases cause the more climate-conscious to be more likely to update NG beliefs downward. However, the climate framing has no significant effect on one's likelihood of updating MG beliefs upward.

Are these differences in partisan directional updating robust to the inclusion of covariates? Tables 3.10 and 3.11 below show the effects of partisanship on downward

NG and upward MG adjustment, respectively, with the inclusion of experimental condition (Model 1) and key covariates (Model 2).

Table 3.10: Downward NG adjustment with covariates

	Model 1	Model 2
(Intercept = Sourced fact)	-0.76 ***	-0.83 *
	(0.00)	(0.04)
Condition = Unsourced fact	0.12	0.13
	(0.65)	(0.61)
Condition = Partisan cue	0.33	0.33
	(0.19)	(0.20)
Condition = Climate cue	0.72 **	0.74 **
	(0.00)	(0.00)
Party ID = Independent	-0.44 *	-0.42 *
	(0.03)	(0.05)
Party ID = Republican Party	-0.72 **	-0.61 *
	(0.00)	(0.01)
Age = 30-50		-0.48 *
		(0.03)
Age = 50-65		-0.51 *
		(0.05)
Age = 65+		-0.46
		(0.12)
Gender = female		0.20
		(0.25)
Serious issue		0.04
		(0.26)
N	633	633
Pseudo R2	0.05	0.07

Notes: For partisanship, the reference category is Democratic Party. For age group, the reference category is 18-30. Gender is coded as: 1 = male, 2 = female. Serious issue reports agreement (0-10) with the statement

Climate change is a serious issue that the government should address. Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table 3.11: Upward MG adjustment with covariates

	Model 1	Model 2
(Intercept = Sourced fact)	0.08	0.16
	(0.67)	(0.67)
Condition = Unsourced fact	0.22	0.24
	(0.34)	(0.31)
Condition = Partisan cue	-0.17	-0.18
	(0.47)	(0.44)
Condition = Climate cue	-0.31	-0.31
	(0.18)	(0.18)
Party ID = Independent	0.16	0.13
	(0.41)	(0.49)
Party ID = Republican Party	0.59 **	0.42
	(0.00)	(0.06)
Age = 30-50		0.19
		(0.38)
Age = 50-65		0.52 *
		(0.03)
Age = 65+		0.59 *
		(0.03)
Gender = female		0.07
		(0.65)
Serious issue		-0.04
		(0.14)
N	633	633
Pseudo R2	0.03	0.05

Notes: For partisanship, the reference category is Democratic Party. For age group, the reference category is 18-30. Gender is coded as: 1 = male, 2 = female. Serious issue reports agreement (0-10) with the statement *Climate change is a serious issue that the government should address.* Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the

probability of the outcome variable, while a negative coefficient suggests a decrease in the probability.
*** p < 0.001; ** p < 0.01; * p < 0.05.

As can be seen, the effects of partisanship on downward NG adjustment are consistent across model specification, such that Democrats are significantly more likely than Republicans to lower their favourability toward NG following treatment. However, the observation that Republicans are more likely to adjust MG beliefs upwards is sensitive to the inclusion of covariates, such that it falls below the threshold of significance (p = 0.06) in Model 2. In both cases, there are strong age-based differences in updating behaviour, such that older voters are less likely to adjust NG beliefs down and more likely to adjust MG beliefs up. This tracks with previous research that shows younger voters are more likely to prioritise renewable energy development and support phasing out fossil fuels (Funk, 2021). Moreover, because of significant gaps in partisanship between age groups, with younger voters increasingly aligned with the Democratic Party (Frey, 2022), it may be the case that some of the previous Republican effect on upward MG adjustment is better captured by age differences in Model 2. To test this, I rerun the original partisan models to include age and an interaction between age and partisanship. The results, which are available in Appendix A.3.4, show that Republicans of all age groups remain significantly less likely to adjust their NG downward; but that, for MG adjustment, there is a significant interaction of age and partisanship, such that older Republicans are significantly more likely to adjust MG beliefs upwards.

3.4.4 Is There Backfire?

Finally, in light of previously summarised debate on backfire effects, I test for the presence of any backfire to the treatments. I evaluate this using the previously defined measure of belief divergence, which classifies individuals binarily based on whether their post-treatment gap between NG and MG beliefs is greater than their pre-treatment gap. On the whole, backfire is rather uncommon but not entirely absent:

roughly 15% of participants diverge their beliefs further following the information treatment. As I find no evidence of any significant differences in the prevalence of backfire based on experimental condition or partisanship, I exclude these results from the main text (they can be found in A.3.5). Conversely, individuals who hold extremely favourable views toward MG originally (classified as expressing a favourability rating of 8 or greater) are significantly more likely to diverge following treatment. This result, shown in Table 3.12 below, echoes research showing backfire is most common for those whose beliefs are strongest.

Table 3.12: Extreme belief and backfire

	Backfire
(Intercept)	-1.80 ***
	(0.00)
MG >8	0.81 *
	(0.02)
N	633
Pseudo R2	0.01

Notes: Extreme favourability is a binary variable that takes the value of 1 for individuals whose pre-treatment MG rating > 8. Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

3.4 Discussion

The results presented above offer an interesting view into not only whether but also how people update a misperception when presented with disconfirming information. I find no evidence for a common sketching of individuals as fact-resistant, but rather that, on the whole, beliefs tend to be rather responsive to a fact even when it contradicts previous perceptions. What's more, this is true of respondents across the political spectrum, suggesting that partisanship does not bear on responsiveness to facts, *per se*. I take this as good news.

Conversely, though people may be willing to converge a specific belief in line with the evidence, I observe significant (and predictable) heterogeneity in this updating behaviour. Put simply, when presented with the option to update NG beliefs down or MG beliefs up, people are inclined to update in the direction of their broader partisan attitudes. A lay interpretation for this is that Republicans who are generally favourable toward NG but have initial reservations about MG take the treatment as a welcomed opportunity to be favourable to MG. This echoes some past research that shows people react more strongly to *good news* and less strongly to *bad news* as it relates to their partisan and climate attitudes (Sunstein et al, 2016). In this case it is a balance between 'the thing I thought was good is worse than I thought' and 'the thing I thought was bad is better than I thought', with the scale weighing in the direction of one's broader view. In turn, this asymmetry seems to imply that factual responsiveness on its own is not only not a solution to, but in fact can exacerbate, partisan belief polarisation, as people incorporate an updated factual belief in a way that suits their way of thinking. This finding is in line with one of the views put forward in the literature review: partisans may — in some cases — be equally willing to acknowledge an empirical reality, but the attributions they make about the implications of that reality (e.g., is my preferred political party to blame

for rising inflation?) go strongly in the direction of their previously-held biases (Tilley & Hobolt, 2011; Bisgaard, 2015).

As for the treatments themselves, I find evidence in line with past research on the deleterious effects of partisan cues: belief convergence is significantly lower in T3 where reference is made to partisan debates about NG. This corresponds to a roughly 16% lower mean return to the efficacy of evidence. However, though lower than the other treatments, beliefs do still converge to a significant degree ($\mu = 59.8\%$) even in the presence of the partisan cue. Moreover, I observe no significant conditional effects of the partisan cue on directional changes — or in other words, people do not engage in asymmetric updating in that condition significantly more or less than in the others. One implication of this is that politicisation of a topic makes people perceive corresponding information as less credible or less absolute, and thus to rely more heavily on their previous evaluations, yielding less change. This fits with some past research finding partisan cues lower the perceived value of substantive information (Druckman, Peterson & Slothuus, 2013). Another, more speculative, interpretation could be that the mention of partisan squabbling simply leads people to stop engaging with the topic altogether. By this line of reasoning, respondents could find the politicisation of a factual topic unnecessary or unpleasant, lowering the perceived valence of the information, in turn lowering their demand for new information altogether (Golman et al, 2022).

These findings are not without their caveats. For one, while people do hold strong energy policy views, they may not pertain to a specific gas so much as to broader climate and partisan attitudes — and thus the result may not hold in the context of more entrenched beliefs. One shortcoming in my design was not measuring the strength of these previous beliefs. It is also possible that, were there not two options for directionally changing beliefs (e.g., if updating on information required a counter-

attitudinal shift, rather than accommodated for a shift in a preferred direction) convergence with the information presented would have been less likely. Second, pertaining to the insignificant effect of removing the source of the information: it may be the case that people extended some perceptions of source credibility to the information, knowing, as they did from the details provided in the informed consent, that the research was conducted by academics at a familiar institution.[17] Practically, this is not something for which I could correct in future research. Because the information in the unsourced fact condition remained strictly factual, people may have been less likely to engage in motivated scepticism — this is both a feature of the design, but also one way in which the question under consideration differs from other examinations of belief change. Finally, as I alluded to previously, it seems likely that — partisan directional differences notwithstanding — the pattern of all respondents adjusting their MG beliefs toward their NG beliefs more is at least partially attributable to greater familiarity with the term *natural gas* than the term *methane gas*. Referring to the earlier paradigm: if NG beliefs are more likely to have been previously considered, they may also be less receptive to new information (Burdein, Taber & Lodge, 2006). While familiarity is a product of external factors beyond my control, it underscores the importance of controlling for strength of previous belief in follow-up research.

These findings speak to some of the underlying mechanisms of how people change their beliefs, and as such should be of interest to readers and academics across a range of related inquiry. If there are two main takeaways, they are that beliefs are likely more responsive to new evidence than some past research gives them credit for — especially if that information pertains to a factually provable subject-matter — but that people's patterns of belief assimilation are nonetheless coloured by partisan

[17] Not to mention that the consent form explicitly states a policy against deception.

predispositions and a bias toward interpreting information in a way that suits their broader preferences. The present context presents a unique setting in which to examine the long-debated extent of asymmetries in updating behaviour, and seems to lend itself to the view that partisan motivated reasoning is prevalent in how people construct and maintain their political views. In turn, this process of asymmetric assimilation, if not strictly implying more polarisation, is at least not incompatible with patterns of diverging views. This generally echoes past findings on self-serving bias in judgment and decision-making, and particularly tracks with one well-known study that asserts: "the fact that people interpret information in a self-serving manner means that, contrary to a fundamental derivation of Bayesian theory ... giving two parties more information may cause their expectations to diverge" (Babcock et al, 1995, p.1337). Whether the present results reflect violations to Bayesian updating (a point which is by no means agreed upon; Bullock, 2009) is beyond the scope of this paper. But, achieved through Bayesian means or otherwise, my findings suggest that people learn what they want from new information.

3.5 References

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*(6), 1037.

Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, *29*(3), 369-387.

Aronow, P. M., & Miller, B. T. (2016). Policy misperceptions and support for gun control legislation. *The Lancet*, *387*(10015), 223.

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *The American Economic Review*, *85*(5), 1337-1343.

Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political perceptions. *Political Behavior*, *24*, 117-150.

Benegal, S. D., & Scruggs, L. A. (2018). Correcting misinformation about climate change: The impact of partisanship in an experimental setting. *Climatic change*, *148*(1-2), 61-80.

Bisgaard, M. (2015). Bias will find a way: Economic perceptions, attributions of blame, and partisan-motivated reasoning during crisis. *The Journal of Politics, 77*(3), 849-860.

Bolsen, T., & Druckman, J. N. (2018). Do partisanship and politicization undermine the impact of a scientific consensus message about climate change?. *Group Processes & Intergroup Relations*, *21*(3), 389-402.

Bolsen, T., Druckman, J. N., & Cook, F. L. (2015). Citizens', scientists', and policy advisors' beliefs about global warming. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 271-295.

Brenan, M. (2022). Americans' Trust in Media Remains Near Record Low. Gallup. https://news.gallup.com/poll/403166/americans-trust-media-remains-near-record-low.aspx

Bullock, J. G. (2006). Partisanship and the enduring effects of false political information. *Unpublished Manuscript*.

Bullock, J. G. (2009). Partisan bias and the Bayesian ideal in the study of public opinion. *The Journal of Politics*, *71*(3), 1109-1124.

Burdein, I., Lodge, M., & Taber, C. (2006). Experiments on the automaticity of political beliefs and attitudes. *Political Psychology*, *27*(3), 359-371.

Clemence, M. (2022). The economy and inflation tied as the biggest issues for Britain, while concern about immigration also increases. *Ipsos Issue Index*. https://www.ipsos.com/en-uk/economy-and-inflation-tied-biggest-issues-britain-while-concern-about-immigration-also-increases

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, *85*(5), 808.

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, *6*(1), 2053168018822174.

De Vries, C. E., Hobolt, S. B., & Tilley, J. (2018). Facing up to the facts: What causes economic perceptions?. *Electoral Studies*, *51*, 115-122.

Druckman, J. N., & Nelson, K. R. (2003). Framing and deliberation: How citizens' conversations limit elite influence. *American Journal of Political Science, 47*(4), 729-745.

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, *107*(1), 57-79.

Duch, R. M., Palmer, H. D., & Anderson, C. J. (2000). Heterogeneity in perceptions of national economic conditions. *American Journal of Political Science, 44*(4), 635-652.

Duch, R. M., & Stevenson, R. T. (2011). Context and economic expectations: when do voters get it right?. *British Journal of Political Science*, *41*(1), 1-31.

Eagly, A. H., & Chaiken, S. (1993). *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers.

Ecker, U. K., & Ang, L. C. (2019). Political attitudes and the processing of misinformation corrections. *Political Psychology*, *40*(2), 241-260.

EPA. (2016, June 3). *Overview of the Oil and Natural Gas Industry*. US Environmental Protection Agency. https://www.epa.gov/natural-gas-star-program/overview-oil-and-natural-gas-industry#:~:text=Production%3A%20Taking%20raw%20natural%20gas

Evans, J. S. B., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review*, *103*(2).

Faramawy, S., Zaki, T., & Sakr, A. E. (2016). Natural gas origin, composition, and processing: A review. *Journal of Natural Gas Science and Engineering*, *34*, 34-54.

Fiorina, M. P. (1981). *Retrospective voting in American elections*. New Haven, CT: Yale University Press.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888.

Flynn, D. J. (2016, September). The scope and correlates of political misperceptions in the mass public. In *Annual Meeting of the American Political Science Association, Philadelphia*.

Frey, W. H. (2022, November 18). *Midterm exit polls show that young voters drove Democratic resistance to the "red wave."* Brookings. https://www.brookings.edu/research/midterm-exit-polls-show-that-young-voters-drove-democratic-resistance-to-the-red-wave/

Funk, C. (2021, May 26). *Key findings: How Americans' attitudes about climate change differ by generation, party and other factors*. Pew Research Center; Pew Research Center. https://www.pewresearch.org/fact-tank/2021/05/26/key-findings-how-americans-attitudes-about-climate-change-differ-by-generation-party-and-other-factors

Gaines, B. J., Kuklinski, J. H., Quirk, P. J., Peyton, B., & Verkuilen, J. (2007). Same facts, different interpretations: Partisan motivation and opinion on Iraq. *The Journal of Politics*, *69*(4), 957-974.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*(1), 20-29.

Golman, R., Loewenstein, G., Molnar, A., & Saccardo, S. (2022). The demand for, and avoidance of, information. *Management Science*, *68*(9), 6454-6476.

Highton, B., & Kam, C. D. (2011). The long-term dynamics of partisanship and issue orientations. *The Journal of Politics, 73*(1), 202-215.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, *22*, 129-146.

Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *The Journal of Politics, 74*(3), 672-684.

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1420.

Jones, J. (2021, March 19). *5 Facts About the State of the Gender Pay Gap | U.S. Department of Labor Blog*. Blog.dol.gov. https://blog.dol.gov/2021/03/19/5-facts-about-the-state-of-the-gender-pay-gap

Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research*, *14*(2), 147-174.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature climate change*, *2*(10), 732-735.

Kahneman, D. (2011). Thinking, Fast and Slow. *Farrar, Straus and Giroux*.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.*), Heuristics & biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press. http://dx .doi.org/10.1017/CBO9780511808098.004

Knobloch-Westerwick, S., & Meng, J. (2009). Looking the other way: Selective exposure to attitude-consistent and counterattitudinal political information. *Communication Research*, *36*(3), 426-448.

Lacroix, K., Goldberg, M. H., Gustafson, A., Rosenthal, S. A., & Leiserowitz, A. (2021). Different names for "natural gas" influence public perception of it. *Journal of Environmental Psychology*, *77*, 101671.
Levendusky, M. S. (2010). Clearer cues, more consistent voters: A benefit of elite polarization. *Political Behavior*, *32*, 111-131.

Lieberman, M. D., Schreiber, D., & Ochsner, K. N. (2003). Is political cognition like riding a bicycle? How cognitive neuroscience can inform research on political thinking. *Political Psychology, 24*(4), 681-704.

Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinberg, G., & Rosenthal, S. (2015). Climate change in the American mind: March, 2015. *Yale Project on Climate Change Communication (Yale University and George Mason University, New Haven, CT).*

Levendusky, M. (2009). *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. University of Chicago Press.

Levendusky, M. S., & Malhotra, N. (2016). (Mis) perceptions of partisan polarization in the American public. *Public Opinion Quarterly*, *80*(S1), 378-391.

Lodge M., & Taber C. (2000). Three steps toward a theory of motivated political reasoning. In A. Lupia, M. D. McCubbins & S. L. Popkin (Eds.), *Elements of Reason: Cognition, Choice, and the Bounds of Rationality* (pp. 183–213). Cambridge; New York: Cambridge Univ. Press.

Lodge, M., & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology*, *26*(3), 455-482.

Loewenstein, G., O'Donoghue, T., & Bhatia, S. (2015). Modeling the interplay between affect and deliberation. *Decision*, *2*(2), 55-81.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098.

Luhmann, N. (1979). *Trust and Power*. New York: John Wiley and Sons.

Lupia, A. (2013). Communicating science in politicized environments. *Proceedings of the National Academy of Sciences of the United States of America*, 14048-14054.

McCarthy, J. (2019, March 22). *Most Americans Support Reducing Fossil Fuel Use*. Gallup. https://news.gallup.com/poll/248006/americans-support-reducing-fossil-fuel.aspx

Merpert, A., Furman, M., Anauati, M. V., Zommer, L., & Taylor, I. (2018). Is that even checkable? An experimental study in identifying checkable statements in political discourse. *Communication Research Reports*, *35*(1), 48-57.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological Review*, *106*(1), 3.

Metzger, M. J., Hartsell, E. H., & Flanagin, A. J. (2020). Cognitive dissonance or credibility? A comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research*, *47*(1), 3-28.

Mullinix, K. J. (2016). Partisanship and preference formation: Competing motivations, elite polarization, and issue importance. *Political Behavior, 38*(2), 383-411.

Myers, T., Maibach, E. W., Roser-Renouf, C., Anderson, A., Stenhouse, N., & Leiserowitz, A. (2013). Public Perceptions of NASA's Research and Reactions to the Climate. Nasa. Gov Website. *Nasa. Gov Website (January 25, 2013)*.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*(2), 303-330.

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*, 104768.

Pew Research Center. (2018). *Distinguishing between factual and opinion statements in the news*. The Pew Charitable Trusts. http://www.journalism.org/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news/

Pew Research Center. (2021, April 15). *Americans' views of the problems facing the nation*. Pew Research Center - U.S. Politics & Policy. https://www.pewresearch.org/politics/2021/04/15/americans-views-of-the-problems-facing-the-nation/

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, *34*(2), 243-281.

Saad, L. (2021, October 5). *Are Americans Concerned about Global Warming?* Gallup. https://news.gallup.com/poll/355427/americans-concerned-global-warming.aspx

Schulman, J., Cadenasso, S., Lazer, D., Ognyanova, K., Baum, M., Quintana, A., … Qu, H. (2022, November 4). The COVID States Project #93: Most Important Problems Facing the US Before the 2022 Midterm Election. https://doi.org/10.31219/osf.io/hxewc

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*, 127190. http://dx.doi.org/10.1037/0033-295X.84.2 .127

Slothuus, R. (2010). When can political parties lead public opinion? Evidence from a natural experiment. *Political Communication, 27*(2), 158-177.

Stanley, M. L., Henne, P., Yang, B. W., & De Brigard, F. (2020). Resistance to position change, motivated reasoning, and polarization. *Political Behavior*, *42*, 891-913.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(5), 701-717.

Steentjes, K., Pidgeon, N. F., Poortinga, W., Corner, A. J., Arnold, A., Böhm, G., ... & Sonnberger, M. (2017). *European Perceptions of Climate Change (EPCC)*: Topline findings of a survey conducted in four European countries in 2016.

Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2016). How people update beliefs about climate change: Good news and bad news. *Cornell L. Rev.*, *102*, 1431.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755-769.

Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2021). Exposure to arguments and evidence changes partisan attitudes even in the face of countervailing leader cues. Preprint.

Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, *33*(3), 460-480.

Tilley, J., & Hobolt, S. B. (2011). Is the government to blame? An experimental test of how partisanship shapes perceptions of performance and responsibility. *The Journal of Politics*, *73*(2), 316-330.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124-1131.

Tyson, A., Funk, C., & Kennedy, B. (2022, March 1). *Americans Largely Favor U.S. Taking Steps To Become Carbon Neutral by 2050*. Pew Research Center Science & Society. https://www.pewresearch.org/science/2022/03/01/americans-largely-favor-u-s-taking-steps-to-become-carbon-neutral-by-2050/#:~:text=Two%2Dthirds%20(67%25)%20say

United Nations. (2022, October 21). *Climate change the greatest threat the world has ever faced, UN expert warns*. OHCHR. https://www.ohchr.org/en/press-releases/2022/10/climate-change-greatest-threat-world-has-ever-faced-un-expert-warns

Unsworth, K. L., & Fielding, K. S. (2014). It's political: How the salience of one's political identity changes climate change beliefs and policy support. *Global Environmental Change, 27*, 131-137.

Van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2015). The scientific consensus on climate change as a gateway belief: Experimental evidence. *PloS one*, *10*(2), e0118489.

Van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global challenges*, *1*(2), 1600008.

Walter, N., & Salovich, N. A. (2021). Unchecked vs. uncheckable: How opinion-based claims can impede corrections of misinformation. *Mass Communication and Society*, *24*(4), 500-526.

Weeks, B. E., & Garrett, R. K. (2014). Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 US presidential election. *International Journal of Public Opinion Research, 26*(4), 401-422.

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145-158.

Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, *41*, 135-163.

A.3 Appendices

A.3.1 Randomisation checks

Here I consider the similarity of participant attributes across conditions, in order to test for successful randomisation of participants into experimental controls and treatments. As can be seen, for the most part treatments are balanced, but there is a slight oversampling of females into Treatment 3. However, I find no significant differences in treatment results based on gender, suggesting this should not bias results. All models presented in the main text have been run with the inclusion of gender.

Table A3.1 Randomisation checks

	Age	Gender	Partisanship
(Intercept = Sourced fact)	45.28 ***	1.47 ***	5.08 ***
	(0.00)	(0.00)	(0.00)
Unsourced fact	-0.05	-0.03	-0.38
	(0.98)	(0.55)	(0.31)
Partisan cue	-0.88	0.13 *	-0.28
	(0.64)	(0.02)	(0.44)
Climate cue	-0.85	0.02	-0.52
	(0.66)	(0.69)	(0.17)
N	633	633	633
R2	0.00	0.02	0.00

Note: Gender is coded as 1 = male 2 = female. Partisanship is coded on a 1 to 10 scale where 1 = Strong Democrat and 10 = Strong Republican *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

A.3.2 Accurate perceptions across conditions pre-treatment

Given the slight inconsistency in the number of participants per condition who hold no pre-existing gap in NG and MG beliefs, I test for significant differences between experimental treatments. Table A.3.1 below summarises these results, which show no significant differences between conditions.

Table A3.2 Accurate perceptions across conditions pre-treatment

	Model 1
(Intercept = Sourced fact)	-1.00 ***
	(0.00)
Unsourced fact	-0.25
	(0.35)
Partisan cue	-0.32
	(0.23)
Climate cue	-0.35
	(0.19)
N	633
AIC	684.71
BIC	702.51
Pseudo R2	0.01

Note: Results of logistic regression of condition on the binary variable *No pre gap*, which takes the value 0 for subjects who have a pre-treatment gap in NG and MG beliefs and 1 for those who do not *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

A.3.3 Testing for interaction of condition and climate beliefs

Table A.3.3 Belief convergence with condition x serious issue interaction

	Belief convergence
(Intercept = Sourced fact)	0.35
	(0.13)
Unsourced fact	-0.25
	(0.35)
Partisan cue	-0.59 *
	(0.03)
Climate cue	-0.09
	(0.74)
Party ID = Independent	0.11
	(0.21)
Party ID = Republican Party	0.06
	(0.54)
Age = 30-50	0.31 ***
	(0.00)
Age = 50-65	0.31 **
	(0.00)
Age = 65+	0.51 ***
	(0.00)
Female	-0.02
	(0.82)
Serious issue	0.01
	(0.76)
Unsourced fact x serious issue	0.02
	(0.61)
Partisan cue x serious issue	0.05
	(0.16)
Climate cue x serious issue	-0.00

	(0.95)
N	489
R2	0.07

Note: Gender is coded as: 1 = male, 2 = female. Serious issue reports agreement (0-10) with the statement *Climate change is a serious issue that the government should address.* *** p < 0.001; ** p < 0.01; * p < 0.05.

A3.4 NG and MG belief adjustments with partisan x age interaction

Here I present the interaction model for partisanship and age on belief adjustment, mentioned in 3.4.3 are a few main takeaways from these results. First, the effect of partisanship on NG adjustment is not an artifact of age differences, such that Republicans of all ages are significantly less likely than Democrats to adjust NG beliefs downward. Second, there are within-party differences for Democrats, such that younger Democratic voters are more likely than unlikely to update NG beliefs downward. Third, the previously observed effect of being a Republican on upward MG adjustment appears to have been driven in large part by older Republicans in the 50-65 age group, for whom there is a significant interaction of party and age.

Table A.3.4: Belief adjustments with partisan x age interaction

	NG	MG
(Intercept)	0.07	-0.22
	(0.79)	(0.42)
Party ID = Independent	-0.77	0.60
	(0.05)	(0.12)
Party ID = Republican Party	-0.96 *	-0.12
	(0.03)	(0.77)
Age = 30-50	-0.75 *	0.26
	(0.03)	(0.44)
Age = 50-65	-0.83 *	0.18
	(0.04)	(0.65)
Age = 65+	-0.36	0.63
	(0.41)	(0.15)
Independent x 30-50	0.59	-0.85
	(0.24)	(0.08)
Republican x 30-50	0.33	0.87

	(0.57)	(0.11)
Independent x 50-65	0.54	-0.19
	(0.37)	(0.74)
Republican x 50-65	0.49	1.39 *
	(0.44)	(0.02)
Independent x 65+	-0.39	-0.09
	(0.61)	(0.90)
Republican x 65+	-0.08	0.37
	(0.91)	(0.56)
N	633	633
Pseudo R2	0.05	0.07

Notes: For partisanship, the reference category is Democratic Party. For age group, the reference category is 18-30. Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

A3.5 Testing for backfire effects

Here I present the results of checks for differences in the prevalence of backfire effects based on condition or partisanship. As can be seen, there are no significant differences.

Table A.3.5: Backfire by condition

	Backfire
(Intercept = Sourced fact)	-1.92 ***
	(0.00)
Unsourced fact	0.17
	(0.59)
Partisan cue	0.44
	(0.16)
Climate cue	0.12
	(0.72)
N	633
Pseudo R2	0.01

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a

decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table A.3.6: Backfire by partisanship

	Backfire
(Intercept = Democratic Party)	-1.67 ***
	(0.00)
Party ID = Independent	0.05
	(0.84)
Party ID = Republican Party	-0.23
	(0.40)
N	633
Pseudo R2	0.00

Notes: Coefficients report log-odds ratios. A positive coefficient for a predictor variable indicates a higher log-odds ratio and suggests an increase in the probability of the outcome variable, while a negative coefficient suggests a decrease in the probability. *** p < 0.001; ** p < 0.01; * p < 0.05.

4

4 Does Candidate Identity Affect Voter Response to a Scandal?

Do qualified candidates for public office compete on a level playing field? This paper interrogates whether and how identity-based prejudice might bear on voters' evaluations of candidates in low-information elections. It takes as its starting point evidence showing overt expressions and self-reported levels of prejudice fell considerably in the last half-century, yet subtler forms of discrimination persist. Research in social psychology and behavioural economics suggests some of the reduction in prejudice and discriminatory behaviour is the result of stronger norms against these behaviours, rather than a complete change in individual preferences. In this view, individuals motivated (or simply unconsciously inclined) to discriminate may do so when they can avoid negative stigma by attributing their behaviour to other causes. Simultaneously, the political science literature has documented a change in the treatment of candidates: for example, whereas women candidates were once less popular their male counterparts, this gender gap has largely disappeared in recent years. Do these electoral improvements extend to non-white candidates in the US as well? Moreover, are they the result of changed preferences, or stronger norms against discrimination? Due to their unambiguously negative valence, political scandals provide an ideal context in which to examine the question: by providing an explanation for negative evaluations, individuals may attribute discrimination to ethical concerns. To test this, I design and administer a vignette experiment with eight experimental conditions to measure whether there is a difference in the propensity for candidate evaluations to be discriminatory under different normative conditions. Results suggest that, while there is no overt discrimination on the basis of either race or gender, Black candidates and in particular Black male candidates are punished more harshly for an ethical scandal than are their white counterparts.

4.1 Introduction

Between 1930 and 1970, three influential social psychology studies (dubbed the *Princeton Trilogy*) sought to understand the basis and prevalence of identity-based stereotypes and prejudice in American society. Therein, students at Princeton University were shown a list of 84 adjectives and asked to choose five they believed best characterised members of various ethnic and national groups in the US. The findings revealed that not only were the content of stereotypes against members of ethnic minority groups often quite negative[18], there was also a high degree of consensus about their accuracy. That is, participants repeatedly chose the same characteristics to describe members of these groups. Writing up their results, the authors noted the consensus "seems too great to be the sole result of the students' contacts with members of these races" (Katz & Braly, 1933). Over subsequent decades, a pair of follow-up studies replicated the original results — and while they noted changes in the content of prevailing stereotypes over time, the level of consensus regarding their accuracy remained relatively stable (Gilbert 1951; Karlins et al., 1969).[19] Stereotypes were widespread, and people were not shy to express belief in them.

Yet, in line with considerable changes in race and gender relations in American society since the latter-half of the 20th Century, recent findings suggest progress: not only has the level of consensus on group characteristics faded, so too have the attributes ascribed to minority groups become generally more positive. Over the same period, compelling evidence suggests the perceived gender differences in competence (i.e., "competence equality") have faded almost entirely, despite persistence of perceptions about different competencies (Eagly et al, 2020). Taken together, this

[18] For example, 75% of respondents labelled African Americans 'lazy', 38% 'ignorant', and 22% 'stupid', while the most frequent adjective chosen to describe Turkish Americans was 'cruel' (54%). By contrast, the most common descriptors of English-Americans were 'sportsmanlike' (53%) and 'intelligent' (46%).
[19] Interestingly, the second study observed the lowest level of consensus of the three, before reverting back to the original level in 1969.

research seems to indicate that 'negative stereotypes are consistently fading' (Dovidio & Gaertner, 1991). However, Devine and Elliot (1995) highlight a critical distinction: it is not knowledge of stereotypes that has fallen — rather, it is the extent of personal belief in those stereotypes. Thus, though individuals may be less likely to believe negative stereotypes, they are still very likely to be aware of them.

While, to be sure, awareness of negative stereotypes is preferable to actual belief therein, a host of research on cognitive heuristics suggests that such latent knowledge still affects our evaluations. This is because of the predictive nature of our cognition: our judgment of future events (i.e., *prospection*) is shaped by a combination of our relevant past experiences and a range of contextual factors (Gilbert & Wilson, 2007). Because we cannot have had every experience upon which to base our expectations, we rely on heuristics and implicit associations to simplify our belief formation. These implicit associations are acquired via socialisation, such as knowledge of stereotypes — or what Hinton (2017) calls "culture in the mind". As such, even those with explicit motivations not to act on prejudicial beliefs often fall prey to applying them when making heuristic judgments (Lai et al, 2016).

Thus, even as Americans became less likely to openly express prejudiced beliefs in survey prompts or other overt manners (and, in many cases, less likely to even consciously hold such beliefs), identity-based discrimination remained present in their interpersonal relations. Using unobtrusive measures like pro-social and helping behaviour, non-verbal communication, and aggression, Crosby and co-authors documented the emergence of what they deemed *subtle racism* in the US. They write that, when one examines research using such measures of prejudice, "a less happy conclusion emerges: Antiblack prejudice is still strong among American whites" (Crosby, Bromley & Saxe, 1980, p.560).

These findings gave way to new modes of understanding prejudice: whereas the

former iteration had focused on overt and conscious negative beliefs or affect, the new wave of research concerned *subtle* or *aversive* styles of "attributionally ambiguous" racial discrimination (Reid & Foels, 2010, p. 292). Methods of detecting and differentiating explicit and implicit prejudice were developed (see, for example, Greenwald & Banaji, 1995), the latter showing to persist in Americans' views of minorities (Dovidio & Gaertner, 1998). This new understanding betrayed the unconscious, often automatic, ways in which prejudice was applied, demonstrating that "the negative feelings and beliefs that underlie contemporary forms of bias may be rooted in normal, often adaptive, psychological processes" (Dovidio, 2001, p.834)

A few broad questions emerge from these trends: first, what explains the divergence between overt expressions of prejudice and observed instances of discriminatory behaviour, and is this still present today? Second, how might the legacy of learned stereotypes — even if individuals do not consciously believe them — affect judgments of politicians who are members of the relevant group, and what is the role of social norms in the application of stereotypes? These questions are examined in the context of evaluations of candidates for public office. The goal is to understand whether political candidates compete on a level playing field, and if not, how identity bears on electoral prospects. In the sections that follow, I detail research on the psychological mechanisms of stereotyping, prejudice, and discrimination, as well as literatures on prejudice and political behaviour. On the basis of this past scholarship and analysis of a novel descriptive dataset, I develop theoretical expectations for the subsequent experiment on whether and how gender- and race-based prejudice affects voters' evaluations of candidates in low-information elections. Results indicate the persistence of covert, negative racial discrimination in individuals' perceptions of political candidates, even as overt discrimination and gender-based discrimination amongst white candidates appears to have dissipated.

4.2 Literature Review

4.2.1 Stereotypes, Prejudice, and Discrimination

Though the terms prejudice and discrimination are often used interchangeably in everyday contexts, and are indeed interrelated, there are important distinctions there between. These differences are well explained by the psychologist Susan Fiske, who writes that stereotyping can be thought of as the most *cognitive*, prejudice the most *affective*, and discrimination the most *behavioural* component of category-based response (Fiske, 1988). Like other cognitive heuristics, stereotypes are essentially shortcuts used by individuals to simplify the information around them. They may or may not carry any normatively *good* or *bad* information, and indeed in many cases can be quite adaptive (Gigerenzer, 2008). Put simply, stereotypes are used to inform individuals' expectations about the likely attributes of a member of a given category.

Conversely, prejudice necessarily entails some positive or negative affective or emotional response (by definition one cannot have neutral prejudice). As Dovidio (2001, p.829) writes, "[p]rejudice is commonly defined as an unfair negative attitude toward a social group or a person perceived to be a member of that group." Thus, in the context of political behaviour, stereotypes correspond to an individual's beliefs, while prejudice corresponds to attitudes, which include 'some degree of favour or disfavour' (Eagly & Chaiken, 1993). By this definition, an individual can be aware of a negative stereotype without holding prejudice if that stereotype does not colour their affective response to an individual from the stereotyped group. Similarly, just as stereotypes feed into prejudice, so too does prejudice feed into discrimination. One can hold prejudiced feelings about an individual without acting on this prejudice, in which case there has been no discrimination — whereas if one's behaviour toward another individual is affected by an identity-based prejudice, that is by definition discrimination.

Kunda and Spencer (2003) detail the mechanisms that explain the conditions

under which individuals will be more likely to engage in stereotype-based prejudice. The authors delineate between *stereotype activation* (i.e., "the extent to which a stereotype is accessible in one's mind") and *stereotype application* (i.e., "the extent to which one uses a stereotype to judge a member of the stereotyped group") (Kunda & Spencer, 2003, p.522). As before, the former can be thought of as an automatically occurring heuristic response, while the latter is synonymous with prejudice. The authors describe three 'goals' that influence whether an individual will apply an activated stereotype. These are:

1) *Comprehension goals*, or the need to make sense of events around us;
2) *Self-enhancement goals*, or the need to maintain one's self-identity;
3) *Prejudice-avoidance goals*, or the intrinsic or extrinsic motivations against prejudice.

While comprehension goals may make one more likely to form prejudicial opinions due to the need to simplify their decision environments, prejudice-avoidance goals (obviously) work in the opposite direction. The former enables individuals to quickly categorise information around them, and make decisions on the basis of these automatic processes, sometimes referred to as *System 1* processing (see, for example, Kahneman, 2011). Conversely, prejudice-avoidance goals would be more aligned with the *System 2* mode of thinking, or the "controlled processing" of information (Stanovich & West, 2000, p.658). Whereas much past political science research has assumed that identity- and group-based cognition occurs at the "level of conscious deliberation" (p.366) (i.e., in *System 2*), Burdein and co-authors (2006) show that automatic group-based classifications occur "within the first few milliseconds of processing political information" (p.366), and that in turn these automatic processes colour subsequent cognitive (i.e., *System 2*) evaluations. Thus, even an individual who holds strong prejudice-avoidance goals — but who nevertheless has knowledge of common

stereotypes (even ones they do not believe to be true) — may be led to act prejudicially on the basis of the antecedence of comprehension goals.

It is harder to make clear predictions on the basis of self-enhancement goals, as these are by definition specific to an individual's self-identity. For example, an individual who identifies strongly as an egalitarian would be influenced not to apply an activated stereotype (this also falls under the internal component of prejudice-avoidance goals) (Dovidio, 2001). Ultimately, the extent to which self-enhancement goals may moderate the exercise of prejudice may depend on whether prejudice-avoidance goals are internally or externally motivated: whereas an intrinsically motivated egalitarian would suffer a hit to their self-concept if they were to catch themselves thinking prejudicially, those motivated only by external forces against prejudice might be able to maintain their self-identity if their prejudice occurred without social sanction (Plant & Devine, 1998).

Two critical insights that emerge from this framework form the basis of the research that follows. First, for many individuals, the salience of external motivations against prejudice (e.g., strong anti-prejudice norms, reviewed in 4.2.3) is a consistent predictor of when activated stereotypes will be applied. That is, even an individual who holds prejudiced attitudes can be dissuaded from acting on those attitudes if the incentive against doing so is strong enough. Second, with respect to both internal and external prejudice-avoidance goals, the authors write that "feelings and beliefs about certain groups may nevertheless color reactions to members of these groups when these reactions may also be construed in alternative, nonprejudiced ways" (Kunda & Spencer, p.524). Or in other words, individuals motivated (or even unconsciously inclined) to apply a stereotype or to discriminate may do so when such behaviour can be attributed to other causes. As Dovidio (2001, p.835) summarises it, "Because aversive racists consciously endorse egalitarian values, they will not discriminate directly

and openly in ways that can be attributed to racism; however, because of their negative feelings they will discriminate, often unintentionally, when their behavior can be justified on the basis of some factor other than race."

4.2.2 Measures of Prejudice

In light of the role of external forces in people's decisions to conceal their prejudice, and thus not to discriminate, survey measures are often not the right tool for capturing the prevalence or absence of prejudiced beliefs. As noted in the Introduction, the prominence of research claiming to document the elimination of prejudiced beliefs (in particular, anti-Black prejudice) led scholars to seek measures that could unobtrusively observe prejudice (or the absence thereof) in genuine social interactions. One reason for the limitations of survey response on sensitive issues like race is that, in answering surveys, respondents may be inclined to give the answer they think researchers want to hear, rather than a more honest assessment of their beliefs (also known as social desirability bias; Presser & Stinson, 1998). Referring to the previous paradigm, another explanation could be that those with strong egalitarian views might genuinely believe they do not hold prejudice, but nonetheless behave prejudicially in a real-world context. Thus, observing the true extent of prejudice and discrimination, it was posited, necessitated more subtle techniques.

As mentioned above, a widely used paradigm looks for differences in how likely people are to offer help to strangers on the basis of identity and other factors. In one example — based on the famous lost-letter technique (Milgram, Mann & Harter, 1965) — researchers left a completed graduate school application in a public phone booth, complete with a picture of the supposed author, and an envelope that was already pre-stamped and addressed for delivery. Varying the gender, race, and physical attractiveness of the applicant, the authors found significant differences in how likely the application was to be delivered, with both less-attractive and Black applicants

receiving less help than their more-attractive or white counterparts (Benson, Karabenick & Lerner, 1976). However, not all studies employing similar designs find these differences (see, for example, Bickman & Kazman, 1973), and some contend that race alone does not prevent helping behaviour, but that there is an interaction between race and how responsible one is perceived for their predicament (in this case, being drunk) that does make white actors significantly less likely to help a Black victim (Piliavin, Rodin & Piliavin, 1969). In line with the argument of Dovidio (2001) detailed above, one explanation for this may be the extent to which people feel a victim's responsibility for their own situation provides a reasonable justification for the exercise of prejudice.

Whether race bears on propensity to help in an emergency, further explorations of the differences between self-reported and observed prejudice highlight the often very contradicting results these measures can yield. For example, in light of claims (based on survey research) regarding the decline of racial prejudice in the South, researchers developed a so-called *list experiment* as a less-obtrusive means of capturing prejudice in survey response (Kuklinski, Cobb & Gilens, 1997). Therein, one group of respondents are given a list of, say, three items and asked to report how many (but, crucially, not which) of the items make them angry — while, in a treatment condition, a separate group receive the same list with an additional item relating to race.[20] The results are then normalised and compared. Using this measure, the authors find significant expressions of anti-Black prejudice, far exceeding those recorded in standard survey research. Interpreting these results, they write "prejudiced whites, fearing the label 'racist', avoid expressing their true racial attitudes publicly" (p.346). Further examples of how the study of discrimination has been operationalised experimentally

[20] These items are: 1) the federal government increasing the tax on gasoline; 2) professional athletes getting million-dollar contracts; 3) large corporations polluting the environment; 4) a Black family moving in next door (Kuklinski, Cobb & Gilens, 1997, p.328).

are discussed in the following sections, but the lesson is that overt expressions of prejudice tend not to fully capture the true extent thereof — in particular because people fear being sanctioned for such behaviour.

4.2.3 Norms and Behaviour

A number of studies corroborate the effect of social norms on individuals' actions. In the context of pro-social behaviour, for example, the economists Steve Levitt and John List write that "the greater is the social norm against the wealth maximizing choice, or the degree of scrutiny when the wealth-maximizing choice has a social cost, the larger the deviation from that choice" (Levitt & List, 2007, p.157). That is, when an individual's desired behaviour is at odds with a prevailing norm, the individual will deviate from that behaviour rather than violate the norm — particularly when their decision is under examination. Thus, beyond wanting to please survey experimenters, a broader reason for the discrepancy between covert and overt prejudice is likely the change in norms against openly subscribing to identity-based stereotypes. This is the contention of Zitek and Hebl (2007), who write that as social norms against the display of prejudice — and public opinion on its acceptability — have changed, so too have individuals become less likely to express prejudice openly. However, if the removal of an anti-prejudice norm (i.e., an increase in the acceptability of prejudice) results in discrimination, then, one might infer that an individual's *wealth-maximizing* action was to act on his prejudice, and the norm simply provided adequate incentive to deviate from this behaviour.

Drawing on this understanding, Danilov and Saccardo (n.d.)[21] look for this type of *disguised discrimination* in the context of economic dictator games, in which participants are paired with either a member of their own or a member of a different

[21] A version of the study, albeit not yet peer-reviewed, is available in Saccardo (2015).

ethnic group (signalled by the partner's name). In the control condition, which consists of only a standard dictator game, the authors find no evidence of discrimination on the basis of identity. In the treatment, however, participants are asked to tell a 'white lie' to benefit their counterpart. This breach of an ethical norm is intended to ease the social norm against prejudice by providing an alternative cause for such behaviour. The authors find that, while individuals are willing to breach this norm for a member of their ethnic group, they are significantly less likely to do so to benefit an out-group member. Interpreting these results, the authors write that taste-based discrimination[22] "appears when it can be rationalized as compliance to a virtuous norm" (Saccardo, 2015, p.196). Thus, as predicted by Kunda and Spencer, individuals seem more likely to engage in prejudice-based discrimination when they can provide alternative explanations for such behaviours.

In a particularly interesting and elaborate design, Choi and co-authors (2019) test for differences in helping behaviour across 28 German cities on the basis of not only ethnicity and religion but also normative context. Using a highly coordinated field study, they deploy a female confederate who is either: an immigrant wearing a hijab, the same confederate without a hijab, the same actor visibly wearing a cross, or a German. The relevant norms are varied such that, prior to needing assistance, the confederate either conforms to or does not conform to a social norm (in this case, a norm against littering). Without norm-enforcement, they find significant differences in propensity to come to the aid of an immigrant, but only against the confederate who was wearing the hijab (which they take as signalling a penalty for cultural non-integration and divergent values). With the enforcement of the norm, the gap between the hijab-wearing and German confederate decreases significantly (though still not enough to counteract the

[22] Without delving into a full reckoning of the literature on taste-based versus statistical discrimination, the authors provide a clear explanation of how their experiments isolate the former from the latter. For a thorough review of these concepts, see Ewens, Tomlin & Wang (2014).

effect of discrimination altogether). In a similarly staged follow-up (Choi, Poertner & Sambanis, 2021), the authors seek to control for this potential effect of misaligned values. Once again, they find significant discrimination against a hijab-wearing immigrant, but that it is mediated by a signal the confederate holds progressive gender attitudes, which they interpret as reducing the extent to which a Muslim woman would be punished for perceived conformity to a negative stereotype (i.e., regressive social attitudes). Conversely, no significant effect is observed for native actors based on expression of values (in fact, the native-regressive values condition is the most likely to receive assistance, though the difference with other native conditions is insignificant).

Such observed differences in the extent to which normative contexts, secondary explanations for behaviour, and perceived transgressions affect the incidence of prejudice are the impetus for the experiment that follows. More generally, the logic of these mechanisms aligns with literature on motivated reasoning, whereby individuals will fulfil their *directional goals* to the extent that they can provide a justification for so doing (Kunda, 1990). If a prejudiced individual has a directional goal of acting on their prejudice (i.e., discriminating), the presence of an explanation that minimises focus on prejudice — providing other plausible explanations for the behaviour (e.g., *she's a litterer*) — could provide enough *justification* for them to do so. This understanding of how prejudice (conscious or not) can lead to discrimination underlies the present research.

4.2.4 Does Prejudice Affect Voting Behaviour?

A growing strain of literature in political science seeks to investigate whether and how prejudice affects the evaluation of politicians (albeit mostly by gender). The results of this line of research are mixed: while older research on voting behaviour in Western democracies posited that women might face discrimination in pursuing public office (see, for example, Kelley & McAllister, 1984), the weight of evidence in recent years has

largely shifted against this perspective. For example, in a meta-analysis of 67 studies on candidate choice, Schwarz and Coppock (2022) find that not only do women not suffer at the polls based on their gender, but in fact now enjoy a slight electoral advantage over their male counterparts. Similarly, a study of some 21,000 unique candidates in Canada found that, though a gender gap of roughly 2.5 percentage points existed in the mid-1900s (i.e., in favour of men), women and men now face statistically indistinguishable odds in their bids for office (Sevi, Arel-Bundock & Blais, 2019). As for media coverage of men and women's campaigns, Hayes and Lawless (2015) find no significant differences in their review of portrayals of male and female candidates in local media. Considering representation, whereas women represented just 13.5% of all US congressional districts in 2002 (Huddy & Capelos, 2002), that number has more than doubled to 29% in the 118th Congress (Leppert & DeSilver, 2023). Moreover, though voters (and low-information voters in particular) may still rely on gender stereotypes in evaluating female candidates, gender stereotypes do not "prevent women candidates from attaining political office once they are on the ballot" (Huddy & Capelos, 2002, p.49). Thus, the state of the literature seems to suggest that — whether or not significant gender gaps existed in the past — women do not face outright discrimination from voters in the present.

Nonetheless, there are other ways in which candidates' gender might bear on their electoral prospects. For example, Mo (2015) finds that female candidates are expected to be more qualified before running for office than are their male counterparts — and also that differences in implicit and explicit attitudes toward women affect how candidate-specific information is treated. Dolan (2014) writes that even though female candidates no longer face 'overt hostility' from voters, perceptions of certain capabilities are still influenced by the content of group stereotypes. While Teele and co-authors (2018) find no outright discrimination against female candidates,

they write that social expectations for candidates (e.g., to be married and have children) lead to a "double bind", as raising children more negatively affects women's career development. Anzia and Bernhard (2022) note the important variation in the application of stereotypes, and call for a more nuanced approach to studying this variation. Making use of real voting data, they show that gender stereotypes hurt women in executive (compared to legislative) races and where constituencies are more conservative, but help when the salient issue is education.

Ono and Burden (2019) corroborate this finding that female candidates are punished in executive, but not legislative, elections — but write that partisanship can steel against this, as individuals are less likely to discriminate when the female candidate is a member of their own party. In a similar comparison of executive versus legislative prospects, Smith and co-authors (2007) found that, when presented with identical biographies of perspective presidential candidates (i.e., an executive election), respondents more negatively evaluated that of the female than the male candidate. This focus on context suggests the extent to which stereotypes are applied depends first on how they are activated, and also how relevant they seem (i.e., if they are believed to advance comprehension).

These results seem to track with the notion that expectations play a crucial role in discrimination. Social norms surrounding traditional societal roles are pervasive and internalised (Hinton, 2017), and can shape perceptions even in the absence of explicit prejudice. For example, while male candidates are rewarded for their ambition, research shows that perceptions of women as power-seeking bear negatively on their electoral prospects (Okimoto & Brescoll, 2010). Moreover, even when individuals might not themselves have a preference to discriminate, knowledge of prejudice in society writ large can affect their beliefs and behaviours. This is what Bateson (2020) dubs *strategic discrimination*: selecting candidates based on our perceptions of their broad electability, in

light of known racial and gender-based prejudice. Examining this perspective, she finds that Americans across the political spectrum consider white males to be more electable (based on notions of broad appeal) than both white and Black women of the same credentials. This research highlights another means through which discrimination can persist beyond individual prejudice: when individuals perceive prejudice amongst their peers, their behaviours, particularly voting behaviours, may implicitly incorporate this prejudice, even when they themselves purport not to hold it.

With respect to candidate race, there is considerably less scholarship on which to form an expectation. However, the research that does exist seems to suggest that racial discrimination persists in voter evaluations of candidates. Canonically, Terkildsen (1993) showed that white respondents evaluated Black candidates more negatively than a white candidate with the same biography. Berinsky and co-authors (2011) found that, when told about a (fabricated) sex scandal regarding either President Obama or Senator John Edwards, white voters punished Obama more than Edwards in their subsequent evaluations. More recently, in her evaluation of strategic discrimination, Bateson (2020) finds strong intersectional effects of race and gender on candidate evaluations, such that white males are the most preferred and Black females the least. This tracks with recent field evidence showing that non-white delegates at the Illinois Republican Convention received roughly 9% fewer votes than their white counterparts, controlling for relevant characteristics (Broockman & Soltas, 2020).

Thus, it seems that identity still plays a role in voters' evaluations of candidates — but that the relationship is not as straightforward as previously assumed. More research is needed to disentangle the conditions under which voters may sanction candidates based on identity, particularly with respect to the interaction of gender and race.

4.2.5 Transgressions and the Activation of Prejudice

Given the strong effect of normative contexts on behaviour, it is important to consider whether and how the results reviewed above might be dynamic to shifts in context. In particular, political scandals (and other perceived transgressions) might conceivably affect the prevalence of prejudice in two ways: first, by providing an explanation through which individuals motivated to discriminate but constrained by social sanction can reasonably justify their discrimination (i.e., *disguised discrimination*) (Dovidio, 2001); second, by stirring automatic negative associations in individuals not consciously motivated to discriminate. This is why the unambiguously negative valence of a scandal is an ideal context in which to consider the potentially uneven standards of behaviour individuals face based on identity: by providing a setting in which an underlying behaviour justifies negative responses, it is possible to observe whether the severity of blowback is constant across groups.

In that vein, recent research on discrimination (particularly gender-based discrimination) often considers responses under multiple normative conditions. For example, Kennedy and co-authors (2016) explore how managers are scrutinised and punished in the face of two types of organisational errors: an accidental mistake (i.e., competence-oriented) and an ethical violation. They find that, though women and men leaders are treated the same for accidents, women managers are punished more severely than their male counterparts for ethical violations. Both findings, they contend, are explained by relevant prescriptive norms around gender: women are not expected to be more competent, and thus face no greater backlash for incompetence; on the other hand, women face an "intensified prescription to be ethical" (p.16), and thus are punished more severely for unethical behaviour. A final study tests these lab-based results in the field with archival data on violations to the American Bar Association's (ABA) professional conduct guidelines for attorneys, finding that female lawyers are

punished more harshly than their male counterparts for similar ethical infractions Kennedy, McDonnell & Stephens, 2016).

Subsequent research has validated this general paradigm: for example, in the wake of a product failure, individuals reported a lower likelihood of purchasing a product from a female-led (versus male-led) company if that failure resulted from an ethical transgression, but not if it resulted from incompetence (Montgomery & Cowen, 2020).[23] Similarly, studying negotiations, Amanatullah and Tinsley (2013) find women are more likely to be sanctioned for self-advocacy than their male counterparts, which they too attribute to different prescriptive norms around agency and ambition. Such differences in *punishment severity* (Kennedy et al, 2016) have been observed in the political behaviour literature as well. For example, Barnes and coauthors (2018) show female politicians are more likely to be punished for sex scandals than their male counterparts, particularly by voters classified as *hostile sexists*. However, using a methodology similar to the one laid out in the sections below, Eggers and coauthors (2018) find no difference in how male and female politicians are punished by voters for corruption scandals — though, interestingly, the observe that female voters are more likely to punish female politicians for corruption than they are to punish male politicians.[24]

These latter studies, build on a now robust literature in economics known as correspondence or audit studies, wherein identical job applications are sent to employers with only, variously, the race or gender of the applicant changed, allowing for attribution of differences in response rates (see, for example, Bertrand and

[23] The authors describe the two types of product failures as such: "In the competence failure version of the article, the company is reported to be previously unaware of defect; in the ethical failure version, the company is said to have known about the defect for 6 months but not to have addressed it because of the high cost of replacing the component" (Montgomery & Cowen, 2020, p.643).
[24] Importantly, their study focuses on already elected incumbents, rather than office-hopefuls.

Mullainathan, 2004).[25] A recent meta-analysis of the (many) studies to employ such a technique in the US finds strong presence of labour market discrimination against Black Americans, and that this level of discrimination has not declined meaningfully since 1989 (Quillian et al, 2017). In a similar design, Pager (2003) finds that black job applicants are punished more severely for having a criminal record than are their white counterparts. This places Black job applicants with a criminal record in something of a conundrum, as research also suggests that, in the absence of a criminal background check, employers may still discriminate against Black applicants on the basis of their perceived likelihood that the applicant has a criminal record (Holzer, Raphael & Stoll, 2006). In both cases, negative stereotypes around black males and criminality (see, for example, Blair, Judd & Chapleau, 2004) have a role: in the former, Black candidates are punished for conforming to a negative stereotype; in the latter, Black candidates are punished for not explicitly deviating from that stereotype.

The mechanisms of these types of discrimination, while difficult to fully parse, are also the subject of considerable investigation. After observing that faculty members were 26% more likely to grant a meeting to white male students than their white female and minority peers if the meeting was requested in a week's time — but that this gap all-but disappeared for meetings on the same-day — Milkman and co-authors (2012) posit an explanation rooted in construal-level theory: because future events rely on abstract construals (i.e., are perceived as less tangible than immediate-term events; Trope & Liberman, 2003), considerations of future events rely more heavily on stereotype-based thinking, a phenomenon they term the *temporal discrimination effect* (Milkman et al, 2012, p.713). This corresponds to general research on how heuristics

[25] A number of follow-up studies have advanced our understanding of the effect originally identified in Bertrand and Mullainathan's *Greg vs. Jamal* study. For a review of how both racial and socioeconomic factors might have affected the original response rates, see Uri Simonsohn's essay (http://datacolada.org/51).

work more broadly: humans "must think with the aid of categories" (Allport, 1958, p.19), and thus stereotypes are often activated in moments of spontaneity (Dovidio & Gaertner, 1998). Thus, the more a context allows for heuristic-based thinking, the more likely it may be to yield judgments rooted in stereotypes. This understanding was echoed in a recent study of anti-immigrant prejudice in Germany, wherein drivers with foreign-sounding names were significantly less likely to be requested for trips on a ride-sharing app when name was the only information provided — but this effect disappeared when names were paired with drivers' ratings from past users. Interpreting their finding, the authors write that "stereotypes regarding particular ethnic groups become more salient and active when other information that could signal trust is scarce" (Tjaden, Schwemmer & Khadjavi, 2018, p.428). Or, to consider the inverse, stereotypes may become more salient, and thus discrimination more common, in the presence of a signal an individual is *not* trustworthy.

Taken together, the sum of this research suggests that — in a number of professional and political contexts — individuals face different standards for what constitutes allowable behaviour. Specifically, tolerance for transgressions seems to depend on both the identity of the transgressor (and the observer) as well as the relevant social norms that the transgression breaks (e.g., ethical versus competence). Thus, even in the absence of outright discrimination, women and minorities can face different standards of evaluation than do their white male counterparts. In turn, these different standards are internalised, and can even affect ambition to assume leadership roles in anticipation of uneven scrutiny (Fisk & Overton, 2019).

4.2.6 Research Puzzle: Prejudice as Political Belief

From the theoretical and empirical evidence summarised above, a few lessons are clear. First, people are categorical thinkers and often rely on heuristics, such as stereotypes, when making evaluations (Gigerenzer, 2008). This is particularly the case with familiar

sociopolitical concepts, which may also prompt automatic affective responses that colour cognition (Taber & Lodge, 2001; Lodge & Taber, 2005). Second, even as more deliberative processes may steel against the application of negative stereotypes, activated stereotypes "often become the lens through which targets of stereotypes are perceived" (Spencer et al, 1998, p.1150). Third, social norms materially affect the incidence of prejudiced behaviour (i.e., discrimination), such that norms against prejudice — like those that have developed in American society over the last half-century — pose a high enough sanction to dissuade even individuals who are inclined toward prejudice against acting on that inclination when their actions are under scrutiny by others (Plant & Devine, 1998; Levitt & List, 2007). In their canonical work *Some Roots of Prejudice*, Allport and Kramer describe this fear of self or social sanction (which rests, they contend, on aversion to shame) as an alleviator of prejudice and "a step toward emancipation from bigotry" (Allport & Kramer, 1946, p.31).

Have modern American politics been emancipated from bigotry, then? The evidence reviewed so far gives mixed expectations as to whether gender and race-based prejudices significantly affect individuals' political beliefs and behaviours, and if so, to what extent. On the one hand, the received wisdom from the political science literature seems to be that women's political prospects are now about equal to men's — and there is a broad consensus that racial dynamics in the US have greatly improved in recent decades. As the dominant racial paradigm has shifted "from inegalitarian to egalitarian" (Mendelberg, 2001, p.24), people have become less likely to express (and perhaps also to hold) negative prejudices on the basis of race. This change has been corroborated in population-based panel survey research, which finds declining levels of anti-Black and anti-Hispanic prejudice over the decade from 2008 to 2018 (Hopkins & Washington, 2020).

Yet, on the basis of the sociopsychological research reviewed, it seems likely to

be the case that — as a result of these changing norms — there are differences between the levels of prejudice observed in surveys and the prejudice that would be observed if the norm were altered. Indeed, some experimental evidence suggests that exposure to prejudiced speech by political leaders increases respondents' own expressed prejudice, a result attributed to a relaxing of the norm against expressing prejudice (Schaffner, 2020; Newman et al, 2020). In design that follows, I conceptualise this difference between *expressed* and *held* prejudice as *latent prejudice*, or prejudice that will be revealed when the normative environment is changed.

With respect to evaluations of candidates for public office, latent prejudice could translate into discrimination not only at the point of voting (where it could nonetheless show given the privacy, and thus lack of scrutiny, for voting decisions) but also in actions like campaign contributions, activism, and even those as simple as answering a poll favourably. Such discrimination could take the form of holding candidates to different standards of behaviour, or displaying less or more tolerance for individuals' transgressions (e.g., in-office political scandals) on the basis of identity. If this were the case, it would mean that candidates for elected office do not compete on a level playing field, but rather, that facets of candidate identity materially impact their prospects of attaining electoral success.

The presence or absence of either type of discrimination, then, is the primary focus of this paper. Specifically, the research questions addressed are:

> *RQ1*: Does identity-based prejudice affect the evaluation of otherwise identical candidate biographies?
> *RQ2*: Does the presence of a scandal affect the incidence of prejudice in candidate evaluation on the basis of identity?
> *RQ3*: Do these measures vary based on respondents' individual characteristics?

The past research reviewed to this point informs the hypotheses that will be tested in the section that follows. On the basis of recent research showing little difference in the electoral chances of women and men candidates (see, for example, Schwarz & Coppock, 2022), I posit:

> *H1:* No overt gender-based discrimination in candidate evaluations

Conversely, given evidence suggesting that Black candidates do suffer electoral penalties on the basis of their identity (Broockman & Soltas, 2020):

> *H2:* Some overt race-based discrimination in candidate evaluations

Finally, following on the understanding elaborated above regarding the role of norms and self-justification in the revelation or suppression of prejudice:

> *H3:* Candidates are punished more for transgression on the basis of a minority identity
>
> *H4:* Revealed discrimination is greater than overt discrimination (i.e., there is latent discrimination)
>
> *H5:* Differences in latent discrimination based on respondent individual characteristics (e.g., women punishing female candidates more severely)

In examining these hypothesis, this paper adds to a rich body of past research by empirically testing for the presence of covert discrimination in the evaluation of candidates. Varying evaluative standards for political candidates would suggest that individuals' predispositions and prejudices substantively affect the way they evaluate political information, events, and actors. This study is the first, to my knowledge, to apply this operationalisation, enabling comparison of overt versus revealed discrimination to evaluations of candidates on the basis of both gender and race.

4.3 Political Scandals in the Wild

4.3.1 Constructing an Observational Dataset

The present research evaluates whether race- and gender-based discrimination is present in candidate evaluations, and if and how a transgression affects responses. To form practical expectations, I first turn to real-world data on political scandals in the US. To do so, I compile a unique dataset matching actual scandals involving members of the US Congress with four other public records of legislators' congressional records and electoral outcomes.

To begin, I make use of a list of political scandals in the US from 1990 to the present compiled and made publicly available by Byrner (2014). The data selection proceeds as follows: the names of individuals involved in these scandals are first checked against a pair of datasets on congressional records and individual legislative effectiveness scores (LES) in the US House and Senate, compiled by Volden and Wiseman (2018) and made public by the Center for Effective Lawmaking at the University of Virginia and Vanderbilt University.[26] I retain scandals that involve individuals whose legislative records are in these sets (i.e., only current and former Members of Congress), and further involve an individual who was an active Member at the time of the event (i.e., excluding scandals involving individuals already removed from office). Based on these records, I am able to account for Members' political party, voting records, seniority, vote share in the previous election, chamber of congress, LES, home state, gender, race, and age.

Next, these data are compared against a comprehensive set of election returns maintained by the Election Data + Science Lab at MIT.[27] Doing so further indicates whether candidates ran for re-election in the cycle following the scandal, and if so their

[26] These data can be found at https://thelawmakers.org/data-download
[27] These data can be found at https://electionlab.mit.edu/data

vote shares in those elections. The latter allows for the construction of the main outcome variable of interest: differences in vote shares in the most recent elections before and after the scandal.[28]

Finally, each individual scandal is independently researched to determine whether the individual in question resigned from office as a result thereof, which is taken as a proxy of the severity. Similarly, to attempt to control for the relative legacy of events, I further check Google search results in the form of ["Candidate name" + "scandal"] and cross-reference these results with Member's publicly available Wikipedia pages, recording whether mention is made of a scandal or ethics concern at the time of the event. On the basis of these searches, I construct a (rather imperfect) metric of relative coverage of various scandals to use as a control in subsequent analyses. The majority of all events (77%) meet this criteria of having received considerable coverage.

4.3.2 Forming Expectations Based on Observational Data

Following this procedure, the remaining dataset (n = 119) contains unique political scandals involving in-office legislators. Unfortunately, likely due to the demographic composition of Congressional office-holders themselves, these data are not at all balanced by identity: only 13% of all scandals involve female Members, while a subsequent 18% and 1% involve Black and Latino Members, respectively. Nonetheless, though numbers in these categories are too small for meaningful statistical evaluation, this allows for the construction of a composite category of scandals involving Members who are not white males (the majority of all observations and, indeed, current and past Members) to proxy the effects of minority identity writ large on the political

[28] Lost votes are calculated only on the basis of general election results to allow for meaningful comparison, as vote shares in contested primaries are often more fragmented than generals. Where candidate elections do not appear in the MIT data in the cycle following the scandal, I consult biographies to verify whether the candidate either 1) resigned, 2) did not resign but did not run for re-election (i.e., retired), or 3) ran for re-election but lost in the primary. I also exclude a handful of cases where incumbents had run uncontested in the previous election cycle, as 100% election shares simply do not occur in contested American elections.

consequences of scandals (n = 38). As the results of these analyses are only meant to guide our expectations for the subsequent experiments, I consider this level of significance reasonable.

Because the main outcome variable, decrease in vote share between elections, does not include candidates who ran for re-election but lost in the primary, I also consider how scandals affect the likelihood a candidate wins re-election (binary yes / no) conditional on having stood for office again (binary yes / no), which expands the case number by roughly a third compared to analyses of changes in vote share. This allows for further study of cases in which candidates lose in a primary following a scandal (quite clearly also a negative outcome), while maintaining the conceptual consistency of the lost vote share measure (i.e., not comparing within-party primary vote shares to general election results).

A first note is that the data are quite evenly split by political party. Moreover, I find no significant differences of candidates' partisanship on voters' responses to scandal, with both Republicans and Democrats losing similar shares of votes (p = 0.62).

Table 4.1: Mean difference between vote share in previous and subsequent election by party

	All	**Republicans**	**Democrats**
Decrease in vote share	9.1%	8.3%	9.8%
N	119	58	61

While I do find suggestive differences in lost votes between male (μ = 8.3%) and female (μ = 13.2%) candidates, there are not enough female observations to perform statistically meaningful tests of independence. Using the composite group of white males (μ = 6.6%) versus all other representatives (μ = 13.9%), however, I observe a significant relationship between identity and declines in vote share following scandals (p

= 0.021).

Probing this further, I estimate a series of linear regression models to see if and how the inclusion of covariates affects the relationship, the results of which are presented in 4.2 below. Model 1 presents only the relationship between identity and lost votes. Model 2 includes a series of regressors to control for Members' electoral records: including their Legislative Effectiveness Score (LES), seniority, and age. Model 3 includes the variable proxying whether scandals are prominently recorded in a Member's biography, while Model 4 further controls for whether candidates are Representatives or Senators. The last inclusion in Model 5 is the type of scandal, which sees the effect of non-white status drop below the threshold of significance.[29] The outcome here is the difference in vote share from the election preceding the scandal and the one following, with a positive value indicating a positive drop (i.e., loss of vote share). As can be seen, minority representatives do suffer a slight penalty in vote share compared to their white male peers, though the significance of this is sensitive the inclusion of covariates.

Table 4.2: Difference in vote share with covariates

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	0.07 ***	-0.08	-0.10	-0.11	-0.07
	(0.00)	(0.38)	(0.29)	(0.25)	(0.57)
Minority = 1	0.07 *	0.08 *	0.08 **	0.08 *	0.06
	(0.02)	(0.01)	(0.01)	(0.02)	(0.11)
Seniority		-0.00	-0.00	-0.00	-0.01
		(0.57)	(0.70)	(0.43)	(0.35)
LES		0.00	-0.00	-0.00	-0.00

[29] There are 11 unique scandal types, but because none of the coefficients are significant, and in the interest of space, I have omitted them from the main text. The full results of this model are available in the A.4.11.

		(0.96)	(0.84)	(0.83)	(0.97)
Age		0.00	0.00	0.00	0.00
		(0.14)	(0.16)	(0.09)	(0.26)
Media coverage			0.03	0.04	0.05
			(0.32)	(0.26)	(0.21)
Senator				-0.05	-0.03
				(0.17)	(0.38)
Abuse of office					0.02
					(0.73)
N	65	65	65	65	65
R2	0.09	0.13	0.14	0.17	0.33

Note: As noted above, the number of cases here is nearly halved from the whole sample of $N = 119$, as it is based only on candidates whose re-election vote tolls are from contested general elections (i.e., not primaries or uncontested generals).
*** p < 0.001; ** p < 0.01; * p < 0.05

I next turn to the question of propensity for re-election conditional on running for office again. Here the results indicate no meaningful impact of identity, irrespective of model specification. These findings are summarised in Table 4.3 below.

Table 4.3: Likelihood of re-election given running again

	Model 1	Model 2	Model3
(Intercept)	0.70 ***	0.98 ***	0.76 ***
	(0.00)	(0.00)	(0.00)
Minority = 1	-0.11	-0.15	-0.13
	(0.35)	(0.16)	(0.30)
Media coverage		-0.37 **	-0.36 **
		(0.00)	(0.01)
LES			-0.01
			(0.68)

Seniority	0.03
	(0.09)
N	-0.00
R2	(0.32)

Note: The number of cases here is slightly greater given inclusion of candidates who do run for re-election but lose in a primary (thus, do not have a vote share). *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

While the coefficient on identity in these models is directionally aligned with the findings above (i.e., negative, or lower likelihood of re-election), we can see that there is no significant effect on re-election, conditional on running again. Thus, though identity seems to impact on incumbents' vote share in the general election, it does not determine their electoral prospects on the whole. Conversely, the variable proxying media coverage of scandals does significantly predict chances of re-election, such that scandals that receive considerable media coverage are, intuitively, more damaging to one's re-election prospects. One explanation for the significance of media coverage here but not in the above models on vote share could be that more severe scandals result in more negative attention, and in turn, this makes politicians more likely to lose in the primary or not run at all (and as such, to not be included in our models on vote share). This squares with findings that, because roughly 60% of all American congressional districts are dominated by a single party, primary contests are often the most effective means of holding elected officials accountable (Hirano & Snyder, 2012).

As for the somewhat unintuitive differences between a significant effect of scandals on subsequent vote share, but not overall likelihood of winning re-election, a pair of explanations stand out. First, it is important to note the inherent selection effects of only considering scandals where the politician stands for re-election: it seems likely that the degree of a scandal bears on whether a candidate is even able to do so, and that the worst scandals (such as those that result in criminal sanctions of the actor

involved) may preclude subsequent runs.[30] Second, past research notes the bind that partisans face in punishing politicians' malfeasance: while, on the one hand, voters want to hold officials accountable, their partisan preferences in the outcome of elections may lead them to be more forgiving of scandals by co-partisans when that official stands for re-election (Eggers, 2014). Thus, partisans may demonstrate their disapproval in the form of de-motivated turnout when they believe that this penalty will not result in their party losing the election, potentially explaining the effect on incumbents' vote share but not overall prospects of winning re-election.

On the whole, the results of these analyses give mixed expectations for the following experiment: while I do find evidence that, in the wake of a scandal, voters may treat candidates differently based on identity, this is sensitive to the inclusion of covariates, and there are too few cases in the dataset to generalise these results or identify specific causes. Nonetheless, there seems to be at least some *a priori* evidence that voters may tolerate transgressions differently based on the identity of the actor.

[30] For example, in the case of Congressman Bill Janklow, who was found guilty of second-degree manslaughter in 2003 less than a year after assuming office, and was thus unable to stand for re-election in 2004.

4.4 Experimental Design & Methods

Following on these observational analyses, I turn to the main empirical strategy of this paper: a between-subjects vignette experiment with eight experimental conditions. I recruit respondents (n=709) from the sample provider Lucid Theorem, which fields US-based subject pools balanced for demographic and political characteristics. Studies evaluating the demographic composition of Lucid samples have found they largely track with nationwide benchmarks, and that the platform is suitable for most types of social science research (Coppock & McClellan, 2019). Participants are then directed to the Qualtrics-based survey and, after passing the conditions of consent, are randomised into an experimental condition.[31]

Therein, participants are shown biographies of hypothetical candidates for state office in Ohio, and asked to indicate their favourability toward the candidate and how likely they would be to vote for them. As Bernhard and Freeder (2018) show, in less prominent races such as local and state office where exposure to information about candidates is lower, voters tend to rely on heuristics and "deal-breakers" to evaluate candidates — and less sophisticated voters often focus on personal information about candidates, rather than policy cues. Thus, the context of a state office election is ideal to parse how identity-based heuristics may affect consideration of political actors, separate from preferences based on policy differences.

The biographies participants review are identical, with only the name of the candidate (indicating race and gender) varied.[32] Each participant sees and evaluates only

[31] In line with the pre-analysis plan, participants who failed to pass a pair of attention and manipulation checks were dropped from the sample, leading to an attrition rate of roughly 10% (original n = 788).
[32] While name differences between gender are often quite obvious, in order to attain a reliable manipulation on race I consult an analysis conducted by Gaddis (2017), which evaluates names used in previous correspondence and audit studies based on what share of respondents perceive that name to be of a given race. I select names from this list for which at least 80% of respondents indicated a clear perception of the race of the individual. After so doing, I consider two additional criteria before making the ultimate selection. First, based on research suggesting that at least some of the differences in past correspondence studies (such as the canonical Bertrand and Mullainathan, 2004) might be attributable to perceptions of low socioeconomic status (SES) related to the Black names used (i.e., Jamal and Lakisha),

one individual candidate biography. In the control conditions, this is the entire study. In the treatments, an additional (brief) article about an investigation into the candidate for unpaid taxes is included in the biographical material.[33] As explained previously, this transgression serves to alter the normative condition, providing another factor to which discrimination can be ascribed. The same measures of preference are then elicited.

The measures of discrimination are based on any observed differences in means related to candidate evaluations between control and treatment groups. The classifications of experimental conditions are as follows:

Controls	**Treatments**
White male, no scandal (C1)	White male, scandal (T1)
White female, no scandal (C2)	White female, scandal (T2)
Black male, no scandal (C3)	Black male, scandal (T3)
Black female, no scandal (C4)	Black female, scandal (T4)

Using the mean likelihood of voting for the candidate and candidate favourability (μ) for each condition (where a higher score indicates a more favourable opinion of the candidate), discrimination is then calculated following the logic:

$$\textit{Overt gender discrimination} = (\mu C1 - \mu C2) \textit{ or } (\mu C3 - \mu C4)$$

$$\textit{Overt race discrimination} = \frac{(\mu C1 + \mu C2)}{2} - \frac{(\mu C3 + \mu C4)}{2}$$

$$\textit{Revealed gender discrimination} = (\mu T1 - \mu T2) \textit{ or } (\mu T3 - \mu T4)$$

I select names that Simonsohn (2016) shows are perceived to be of relatively higher SES. Second, consulting a "partisan name calculator" constructed by Clarity Labs — which checks first names against voting registrations to indicate what percent of Americans with a given name are members of each political party — I choose names that are in proportion to Black and white voters aligned with each party.

[33] The negative treatment is based on an attack article published about then-candidate Jamaal Bowman during his race for the US House of Representatives in New York (https://tinyurl.com/5923ttra).

$$\textit{Revealed race discrimination} = \frac{(\mu T1 + \mu T2)}{2} - \frac{(\mu T3 + \mu T4)}{2}$$

Measures of *latent* discrimination (i.e., the difference in discrimination based on normative contexts) are simply the differences between the overt and revealed measures, for example:

$$\textit{Latent gender discrimination} = (\mu T1 - \mu T2) - (\mu C1 - \mu C2) \textit{ or } (\mu T3 - \mu T4) - (\mu C3 - \mu C4)$$

Recalling the previously stated hypotheses, these measures are used as follows:

H1: No overt gender-based differences (i.e., $(\mu T1 - \mu T2)$ *and* $(\mu T3 - \mu T4) = 0$)

H2: Some overt race-based differences (i.e., $\left[\frac{(\mu C1 + \mu C2)}{2} - \frac{(\mu C3 + \mu C4)}{2}\right] > 0$)

H3: Candidates are punished more for transgression on the basis of a minority identity (i.e., there is revealed discrimination) (e.g., $\frac{(\mu C1 - \mu T1)}{\mu C1} > \frac{(\mu C3 - \mu T3)}{\mu C3}$)

H4: Revealed discrimination is greater than overt discrimination (i.e., there is latent discrimination) (e.g., $(\mu C1 - \mu C3) < (\mu T1 - \mu T3)$)

H5: Differences in latent discrimination based on respondent individual characteristics (e.g., women punishing female candidates more severely)

The main outcome variables for evaluating these hypotheses are respondents' answers to the questions: *Based on the information you've learned from reading the above biography*

1) How favourable are you toward this candidate's campaign for State House? [Likert 0-10]
2) How likely would you be to vote for this candidate for office? [Likert 0-10]

I further collect demographic and political information, notably respondents' age, race, gender, political partisanship, and region, to allow for the consideration of heterogenous responses based on respondent ID.

4.5 Results

4.5.1 Summary of Main Experimental Results

With respect to the research questions stated above, this section analyses the experimental data to consider if, and if so how, candidate identity affects voters' evaluations of their biographies. For reference, measures of overt discrimination refer to mean differences observed in the control conditions (i.e., identical biographies without a scandal); measures of revealed discrimination refer to mean differences observed in the treatment conditions (i.e., identical biographies with a scandal); and measures of latent discrimination refer to the differences between these two. The first two rows in Tables 4.4 and 4.5 below present condition by condition means for the main outcome variables, voting likelihood and favourability, while the final rows show the electoral toll that candidates of each identity suffer for involvement in a scandal (in both absolute and percentage terms).

Table 4.4: Mean candidate favourability by condition

	White Male	***White Female***	***Black Male***	***Black Female***
No scandal	7.10	7.13	7.29	7.12
Scandal	5.21	5.19	4.40	4.98
Penalty	1.89	1.94	2.89	2.15
(%)	26.7%	27.2%	39.6%	30.1%

Note: Favourability reported on a 0-10 scale from very unfavourable to very favourable.

Table 4.5: Mean likelihood of voting for candidate by condition

	White Male	*White Female*	*Black Male*	*Black Female*
No scandal	6.99	6.91	6.91	7.04
Scandal	4.93	5.06	4.19	4.72
Penalty	2.06	1.85	2.71	2.33
Penalty (%)	29.4%	26.8%	39.3%	33.0%

Note: Voting likelihood reported on a 0-10 scale from very unlikely to very likely.

To summarise the main results briefly: first, I find no evidence of overt race or gender discrimination in candidate evaluations. Both of the primary outcome measures are nearly equal across control conditions, and if anything show a very slight preference for Black females in voting likelihood and Black males in favourability. Thus, I find support for *H1* and can reject *H2*; candidates of all four identities are evaluated roughly the same on the basis of their (identical) biographies. Similarly, when considering the entire sample, there is no evidence of revealed gender discrimination in the treatment conditions (though, as detailed below, some groups do punish women more harshly) leading to a partial rejection of *H3*.

Conversely, there is evidence indicating revealed discrimination on the basis of race. In particular, when considering Black male candidates enmeshed in a scandal, participants indicate a 39.3% lower likelihood of voting for that candidate than the comparable Black male candidate without a scandal (for whom participants report a 69.1% voting likelihood). This means that Black male candidates incur a roughly 25% greater electoral penalty than their white male counterparts (who incur a 29.3% penalty) for partaking in the same unethical behaviours. As for women candidates, though the scandal gap is considerably less pronounced on voting likelihood, Black women candidates (33.0%) still incur a roughly 20% greater electoral penalty for the scandal than do their white women (26.4%) counterparts. Thus, I find some evidence in favour

of *H3* as it relates to revealed racial discrimination. Similarly, given that there is no evident gender or race discrimination in the controls, but there is revealed race discrimination, there is also evidence of latent race discrimination in line with *H4*. Recalling the measure of revealed race discrimination, using candidate favourability as an example, I observe a 5.3 percentage point, or roughly 8%, drop in favourability:

$$\frac{(\mu T1 + \mu T2)}{2} - \frac{(\mu T3 + \mu T4)}{2} = \frac{(5.21 + 5.19)}{2} - \frac{(4.40 + 4.98)}{2} = 0.53.$$

Finally, *H5* posited that the extent of latent discrimination would vary based on respondents' characteristics, in particular political party. Accordingly, I find strong evidence that Republicans punish Black candidates enmeshed in scandal more severely than their white counterparts, based on the comparison of evaluations in the treatments to those in the controls. These findings are considered in more depth in section 4.5.3 below.

4.5.2 Examination of Main Experimental Results

In order to test whether any of these differences rise to the level of statistical significance, I estimate four linear regression models for both of the main DVs in the controls and treatments. To do so, I subset the data such that control conditions are only compared with other control conditions and treatment conditions are compared with treatment conditions (i.e., to look for identity-based differences within the same normative contexts). In line with the findings detailed above, I observe no significant differences between control conditions on responses for the main outcome variables — indicating that, across the whole sample, there is no overt discrimination against candidates of any identity without a scandal. Conversely, condition by condition comparisons for the treatments corroborate the gap observed above: Black male candidates (but not Black female candidates) are rated significantly lower in

favourability when the biography includes mention of an ethical scandal. As for voting likelihood, though the gap and direction thereof are similar in magnitude to favourability, the difference does not meet the threshold of significance ($p = 0.08$). These results are summarised in Tables 4.6 and 4.7 below.

Table 4.6: Linear regression for main DVs by condition (controls only)

	Favourability	Vote for
(Intercept = White male, no scandal)	7.09 ***	6.98 ***
	(0.00)	(0.00)
White female, no scandal	0.05	-0.06
	(0.85)	(0.85)
Black male, no scandal	0.21	-0.07
	(0.43)	(0.81)
Black female, no scandal	0.04	0.07
	(0.89)	(0.83)
N	362	362
R2	0.00	0.00

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

Table 4.7: Linear regression for main DVs by condition (treatments only)

	Favourability	Vote for
(Intercept = White male, scandal)	5.22 ***	4.93 ***
	(0.00)	(0.00)
White female, scandal	0.01	0.16
	(0.99)	(0.71)
Black male, scandal	-0.81 *	-0.74
	(0.05)	(0.08)
Black female, scandal	-0.24	-0.21

	(0.56)	(0.62)
N	347	347
R2	0.03	0.02

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

To further examine these results, I construct datasets for condition by condition comparisons within both controls and treatments. This allows me to pit individual identities head to head (e.g., *white male versus white female, no scandal*) and expand the models to include covariates and potential interactions there between (particularly between experimental condition and partisanship). As expected, there is no evidence of overt race or gender discrimination in the controls when considering the full sample. For the most parsimonious models, which compare only responses on the main DVs based on candidate identity (i.e., condition), none of the head-to-head comparison models yield significant results. That is, there is no evidence of overt discrimination against candidates of any identity without a scandal.

When expanded to include covariates and an interaction term between condition and partisanship (i.e., whether responses to specific candidates differ by party), two main results stand out. First, when a female candidate is included in the comparison, gender becomes a significant predictor such that women rate women candidates more highly than their male counterparts. Second, when comparing only Black male and Black female candidates, there is a significant interaction of condition and partisanship for Republicans in Control 4, such that Republicans are significantly less favourable toward Black women than Black men ($p = 0.02$). For voting likelihood, though the gap and direction are similar, the difference ($p = 0.06$) does not meet the threshold of significance. Interestingly, this is the only significant interaction between partisanship and condition across all control models (and in fact, Republicans rate Black male candidates slightly higher than white male candidates without a scandal),

indicating there are no partisan differences between same-gender candidates of the same race, or between white men and white women candidates. The results of the model comparing C3 and C4 are in Table 4.8 below, while the rest of the condition by condition models for the controls are available in A.4.7.

Table 4.8: Linear regression for main DVs, C3 vs C4, with covariates and condition x party interaction

	Favourability	Vote for
(Intercept = Black male, no scandal)	7.20 ***	7.42 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	0.39	0.75
	(0.40)	(0.15)
Age	-0.00	-0.01
	(0.63)	(0.26)
Gender = female	0.83 **	0.83 **
	(0.00)	(0.00)
Region 2 = Midwest	-0.39	-0.53
	(0.34)	(0.25)
Region 3 = South	0.00	-0.12
	(1.00)	(0.76)
Region 4 = West	0.13	0.13
	(0.74)	(0.77)
Party ID = Independent	-0.20	-0.39
	(0.66)	(0.43)
Party ID = Republican	0.37	0.11
	(0.40)	(0.83)
Race = White	-0.23	-0.40
	(0.48)	(0.28)
Black female, no scandal x Independent	-0.24	-0.43
	(0.71)	(0.55)
Black female, no scandal x Republican	-1.48 *	-1.34

	(0.02)	(0.06)
N	184	184
R2	0.11	0.11

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

Whereas the review of conditional mean responses presented above set an expectation that these condition by condition analyses would yield no discrimination in the controls, the significant gaps observed between candidates *with a scandal* sets the opposite expectation for the treatments. To that end, comparisons for the treatments are in line with what was observed previously: in the parsimonious models, which consider variation in responses to DVs based only on condition, only Black male candidates are evaluated significantly lower than their peers. As before, when compared to an identical white male candidate, Black men's favourability and voting likelihood are both roughly 15% lower, though the latter does not rise to the standard of significance ($p = 0.08$). As this is the only head-to-head comparison to register significant differences by condition, I present the results of T1 versus T3 in Table 4.9 below, while the rest of the comparison models can be found in A.4.8.

Table 4.9: Linear regression for main DVs, T1 vs T3

	Favourability	Vote for
(Intercept = White male, scandal)	5.22 ***	4.93 ***
	(0.00)	(0.00)
Condition = Black male, scandal	-0.81 *	-0.74
	(0.05)	(0.08)
N	177	177
R2	0.02	0.02

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

Expanding these models to include covariates and an interaction term for condition and partisanship, a number of interesting findings arise. First, as detailed in the following section, candidate evaluations are strongly affected by the interaction of condition and partisanship. Second, the previously observed difference in favourability between white and Black male candidates enmeshed in scandal is somewhat sensitive to the inclusion of covariates. For example, with respect to the simple model, the inclusion of participant age and gender yields a slight increase in the favourability gap, however including region and partisanship sees this gap fall below the threshold of significance (p = 0.054). Model 3 further includes participant race, which — though race itself does not affect results — yields the greatest estimated significance on the favourability gap (p = 0.026). As shown in Model 4, there is no significant interaction between condition and partisanship in this comparison. In other words, while Black male candidates suffer the greatest scandal gap in favourability, participants of all partisan preference appear to discriminate against Black male candidates at similar levels (relative to white male candidates). These results, slightly truncated to exclude non-significant coefficients on race, are presented in Table 4.10 below, and a full version of all four models is available in A.4.5.

Table 4.10: Tiered covariate regression models for favourability, T1 vs T3

	Model 1	Model 2	Model 3	Model 4
(Intercept = White male, scandal)	5.32 ***	5.33 ***	4.02 *	3.39
	(0.00)	(0.00)	(0.03)	(0.08)
Condition = Black male, scandal	-0.83 *	-0.80	-0.95 *	-0.53
	(0.04)	(0.05)	(0.03)	(0.44)
Age	-0.01	-0.01	-0.01	-0.00
	(0.54)	(0.57)	(0.64)	(0.75)
Gender = female	0.50	0.48	0.67	0.69

	(0.22)	(0.25)	(0.12)	(0.10)
Region = Midwest		0.03	0.03	-0.01
		(0.96)	(0.96)	(0.99)
Region = South		-0.36	-0.29	-0.24
		(0.52)	(0.60)	(0.67)
Region = West		-0.03	0.14	-0.01
		(0.95)	(0.81)	(0.98)
Party ID = Independent		0.10	0.25	0.14
		(0.84)	(0.63)	(0.84)
Party ID = Republican		0.20	0.41	1.16
		(0.70)	(0.44)	(0.11)
Black male, scandal x Independent				0.14
				(0.89)
Black male, scandal x Republican				-1.54
				(0.14)
N	177	177	177	177
R2	0.03	0.04	0.09	0.10

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West.
*** p < 0.001; ** p < 0.01; * p < 0.05.

4.5.3 Heterogenous Treatment Effects

Because of the centrality of identity to these questions of potential prejudice and discrimination, I also explore heterogeneous treatment effects based on respondent characteristics. Interestingly, I find no such differences on the basis of respondent gender (excepting those noted previously), such that women are no less or more likely to punish women candidates in the presence of a scandal. Unfortunately, because there are too few Black respondents in each condition to meaningfully consider potential differences, I am unable to account for the effect of respondent race on scandal gaps. And, while the composite categorisation of respondents as either white or non-white

yields no significant effects, it seems likely the result would be different were there enough Black respondents to make these comparisons. Nonetheless, white respondents are at least no likelier to discriminate against Black politicians than are other non-Black respondents.

Conversely, there are interesting differences in respondents' evaluations — notably, with respect to the size of scandal gaps for varied candidate identities — on the basis of partisanship. These differences can be seen in tables 4.11 and 4.12 below.

Table 4.11: Mean candidate favourability by condition (Republican respondents)

	White Male	***White Female***	***Black Male***	***Black Female***
No scandal	6.83	7.10	7.58	6.47
Scandal	5.81	5.42	3.96	4.10
Penalty	1.03	1.67	3.62	2.37
Penalty (%)	15.0%	23.6%	47.7%	36.6%

Note: Favourability reported on a 0-10 scale from very unfavourable to very favourable.

Table 4.12: Mean candidate favourability by condition (Democratic respondents)

	White Male	***White Female***	***Black Male***	***Black Female***
No scandal	7.21	7.47	7.24	7.75
Scandal	5.09	5.06	4.39	6.46
Penalty	2.13	2.41	2.85	1.29
Penalty (%)	29.5%	32.3%	39.3%	16.6%

Note: Voting likelihood reported on a 0-10 scale from very unlikely to very likely.

In particular, Republicans more heavily punish the Black male candidate for a scandal (that is, more heavily than even Democratic and Independent respondents, for whom the scandal gap for the Black male candidate is also disproportionately large), while showing significantly greater leniency for the white candidates (in particular the white

male candidate). On the other side, Democrats show by far the greatest leniency for the Black female candidate — while conforming to the broader cross-partisan trend of punishing the Black male candidate significantly more punitively for involvement in a scandal than his peers.

One interpretation of these findings underscores the potential confounding effect of perceived partisanship on these evaluations. Consider, for example, the somewhat staggering fact that of the 29 Black women currently serving in the US Congress, all 29 represent the Democratic Party.[34] In light of this, it is not unreasonable to expect that Democratic voters may punish the Black female candidate least for a scandal, as they infer the strongest cue that the candidate in question is an in-partisan; with the inverse holding for Republicans. Similarly, while white representatives are still a majority in both parties, a white member of Congress is roughly twice as likely to be a Republican as a Democrat. In such an understanding, the mechanism of discrimination works similarly for partisanship as it might for race: put simply, we afford more or less leniency toward individuals we perceive to be member of an in- or out-group, respectively. The fact that there are no such partisan differences without a scandal lends credence to the notion that one facet of discrimination is less tolerance for transgressions. A particularly fruitful area of further research, then, could replicate and expand on the current design, adding a further dimension of in- and out-partisanship, allowing to further disentangle the causes of this sort of out-group discrimination.

Despite these partisan differences, it is important to note again that, across party lines, Black male candidates are punished significantly more severely than their peers for a political scandal. That this holds even amongst Democrats strongly indicates that race-based prejudices beyond partisanship are a driving factor.

[34] Data from the Center for American Women and Politics (CAWP) at Rutgers University (link here).

4.6 Discussion

Following a rich tradition of research providing experimental evidence to the debate on racial and gender discrimination in the US, the present study sought to examine the extent to which political actors are evaluated on a level playing field. It takes inspiration from many past examinations that find overt prejudice is unlikely to be expressed, but that subtler forms of prejudice nonetheless exist and can result in discrimination if the norms against such behaviours are altered. To isolate these effects, I make use of and innovate on a common design for measuring if and how identity affects response to otherwise identical individuals, further adding a change in the normative context in the form of a political scandal. Specifically, I compare differences in evaluations of an individual candidate (e.g., white male) when the information provided on that candidate either does not (control) or does (treatment) mention the scandal. The conditions are otherwise identical, presenting a brief, partisan-neutral biography with only the candidate's name varied to indicate gender and race.

In line with much recent research, I observe no gender gap in candidate evaluations, neither in the presence nor absence of such a transgression. Conversely, I find strong evidence of revealed discrimination against Black candidates, and in particular Black male candidates, such that they are penalised significantly more harshly than their white counterparts for the same negative behaviour. Crucially, there is no such discrimination in the controls, with Black and white candidates of both genders evaluated to be roughly equivalent in terms of respondents' favourability toward and likelihood of voting for them, both of which are in the order of 70%. As such, I am able to attribute discrimination observed in the treatments as evidence of latent prejudice — which I previously defined as prejudice that exists but is suppressed for fear of social or self-sanction, and can be revealed in the absence of such concerns. While all candidates are, predictably, punished for the scandal, each receiving

significantly lower evaluations *vis-à-vis* their scandal-free alter egos, Black males suffer by far the largest scandal gap, roughly 25% greater than that incurred for the comparable white male candidate. Two additional results stand out: first, the heterogeneous analyses show the potentially confounding effect of perceived partisanship on tolerance for transgression, and the need for further research to expand on the current paradigm while adding a further dimension for in- and out-partisanship. Second, relatedly, regression results show that partisan interaction is not significant for the consideration of Black male versus white male candidates, indicating that respondents across parties are similarly more severe in their punishments of Black male candidates.

Of course, these results are limited by both the context of an experimental evaluation of prejudice as well as some of the specific choices I made in operationalising that evaluation. I note without reservation that — given the brevity of information provided in the conditions — this paradigm is likely not generalisable to actual election results, particularly in higher-profile races in which information abounds. With that said, I believe these findings do reflect strongly on people's evaluations of political actors (and individuals) generally, in particular as they pertain to the varying standards of tolerance afforded to others on the basis of their identity. As such, this research provides a necessary first, but by no means final, step in conceptualising questions regarding the role of racial prejudice in political behaviours (particularly voting), and it provides lessons for future research that may seek to take these evaluations into the field. It also speaks more broadly to questions about when, how, and why people are likelier to suppress, or express, their prejudices.

My findings highlight the ways in which racial prejudice, however it may compare to that which existed in the past, is still a potent force in American politics. They also underscore the necessity measuring prejudice in ways that are able to account

for factors known to affect people's propensity to show it. In line with the mechanisms set out in the literature review, it appears that even relatively minor instances of unethical behaviour are enough justification for people to act on prejudices they hold, whether consciously or otherwise. Though these results may have foreboding implications for the ways in which the legacy of racial prejudice in America lives on today, they also open the door for conversations about prejudice better informed by evidence. As James Baldwin wrote, with characteristic clarity of thought, "not everything that is faced can be changed; but nothing can be changed until it is faced."

4.7 References

Ahler, D. J., & Sood, G. (2018). The parties in our heads: Misperceptions about party composition and their consequences. *The Journal of Politics*, *80*(3), 964-981.

Amanatullah, E. T., & Tinsley, C. H. (2013). Punishing female negotiators for asserting too much… or not enough: Exploring why advocacy moderates backlash against assertive female negotiators. *Organizational Behavior and Human Decision Processes*, *120*(1), 110-122.

Allport, G. W. (1958). *The Nature of Prejudice* (abridged). Garden City, NY: Doubleday. (Original work published 1954).

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. *The Journal of Psychology*, *22*(1), 9-39.

Anzia, Sarah F., & Bernhard, Rachel. (2022). Gender Stereotyping and the Electoral Success of Women Candidates: New Evidence from Local Elections in the United States. *British Journal Of Political Science, 52*(4), Pp1544-1563.

Bateson, R. (2020). Strategic discrimination. *Perspectives on Politics*, *18*(4), 1068-1087.

Barnes, T. D., & Beaulieu, E. (2019). Women politicians, institutions, and perceptions of corruption. *Comparative Political Studies*, *52*(1), 134-167.

Barnes, T. D., Beaulieu, E., & Saxton, G. W. (2018). Sex and corruption: how sexism shapes voters' responses to scandal. *Politics, Groups, and Identities*, 1-19.

Benson, P. L., Karabenick, S. A., & Lerner, R. M. (1976). Pretty pleases: The effects of physical attractiveness, race, and sex on receiving help. *Journal of Experimental Social Psychology*, *12*(5), 409-415.

Bernhard, R., & Freeder, S. (2018). The More You Know: Voter Heuristics and the Information Search. *Political Behavior*, 1-21.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991-1013.

Bickman, L., & Kamzan, M. (1973). The effect of race and need on helping behavior. *The Journal of Social Psychology*, *89*(1), 73-77.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*(10), 674-679.

Broockman, D. E., & Soltas, E. J. (2020). A natural experiment on discrimination in elections. *Journal of Public Economics*, *188*, 104201.

Burdein, I., Lodge, M., & Taber, C. (2006). Experiments on the automaticity of political beliefs and attitudes. *Political Psychology*, *27*(3), 359-371.

Choi, D. D., Poertner, M., & Sambanis, N. (2019). Parochialism, social norms, and discrimination against immigrants. *Proceedings of the National Academy of Sciences*, *116*(33), 16274-16279.

Choi, D. D., Poertner, M., & Sambanis, N. (2021). The hijab penalty: Feminist backlash to Muslim immigrants. *American Journal of Political Science, 00*(0), 1-16.

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, *87*(3), 546.

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, *6*(1), 2053168018822174.

Cox, D. R. (2009). Randomization in the design of experiments. *International Statistical Review*, *77*(3), 415-429.

Danilov, A., & Saccardo, S. (2016). Disguised Discrimination. *Working Paper*.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and social psychology bulletin*, *21*(11), 1139-1150.

Dolan, K. (2010). The impact of gender stereotyped evaluations on support for women candidates. *Political Behavior*, *32*(1), 69-88.

Dolan, K. (2014). Gender stereotypes, candidate evaluations, and voting for women candidates: what really matters?. *Political Research Quarterly*, *67*(1), 96-107.

Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues*, 57(4), 829-849.

Dovidio, J. F., & Gaertner, S. L. (1991). Changes in the expression and assessment of racial prejudice. In H. J. Knopke, R. J. Norrell, & R. W. Rogers (Eds.), *Opening doors: Perspectives on race relations in contemporary America* (pp. 119-148). University, AL, US: The University of Alabama Press.

Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3–32). Newbury Park, CA: Sage

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist*, 75(3), 301.

Eggers, A. C. (2014). Partisanship and electoral accountability: Evidence from the UK expenses scandal. *Quarterly Journal of Political Science*, *9*(4), 441-472.

Eggers, A. C., Vivyan, N., & Wagner, M. (2018). Corruption, accountability, and gender: Do female politicians face higher standards in public life?. *The Journal of Politics*, *80*(1), 321-326.

Ewens, M., Tomlin, B., & Wang, L. C. (2014). Statistical discrimination or prejudice? A large sample field experiment. *Review of Economics and Statistics, 96*(1), 119-134.

Fisk, S. R., & Overton, J. (2019). Who wants to lead? Anticipated gender discrimination reduces women's leadership ambitions. *Social Psychology Quarterly*, 82(3), 319-332.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 357–411). McGraw-Hill.

Gaddis, S. M. (2017). How black are Lakisha and Jamal? Racial perceptions from names used in correspondence audit studies. *Sociological Science*, *4*, 469-489.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*(1), 20-29.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351-1354.

Gilbert, G. M. (1951). Stereotype persistence and change among college students. *The Journal of Abnormal and Social Psychology*, *46*(2), 245.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychological Review, 102(1), 4.

Hayes, D., & Lawless, J. L. (2015). A non-gendered lens? Media, voters, and female candidates in contemporary congressional elections. *Perspectives on Politics*, *13*(1), 95-118.

Hinton, P. (2017). Implicit stereotypes and the predictive brain: cognition and culture in "biased" person perception. *Palgrave Communications*, *3*(1), 1-9.

Holzer, H. J., Raphael, S., & Stoll, M. A. (2006). Perceived criminality, criminal background checks, and the racial hiring practices of employers. *The Journal of Law and Economics*, *49*(2), 451-480.

Hopkins, D. J., & Washington, S. (2020). The rise of Trump, the fall of prejudice? Tracking white Americans' racial attitudes via a panel survey, 2008–2018. *Public Opinion Quarterly, 84*(1), 119-140.

Huddy, L., & Capelos, T. (2002). Gender stereotyping and candidate evaluation: Good news and bad news for women politicians. *The Social Psychology of Politics*, 29-53.

Kahneman, D. (2011). Thinking, Fast and Slow. *Farrar, Straus and Giroux*.

Karlins, M., Coffman, T. L., & Walters, G. (1969). On the fading of social stereotypes: studies in three generations of college students. *Journal of personality and social psychology*, *13*(1), 1.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*(3), 280.

Kelley, J., & McAllister, I. (1984). Ballot paper cues and the vote in Australia and Britain: alphabetic voting, sex, and title. *Public Opinion Quarterly*, *48*(2), 452-466.

Kennedy, J., McDonnell, M. H., & Stephens, N. (2016). Does gender raise the ethical bar? Exploring the punishment of ethical violations at work. *Exploring the Punishment of Ethical Violations at Work (April 25, 2016). Vanderbilt Owen Graduate School of Management Research Paper*, (2770012).

Kuklinski, J. H., Cobb, M. D., & Gilens, M. (1997). Racial attitudes and the" New South". *The Journal of Politics*, *59*(2), 323-349.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological bulletin*, *129*(4), 522.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., ... & Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001.

Leppert, R., & DeSilver, D. (2023, January 3). *118th Congress has a record number of women*. Pew Research Center. https://www.pewresearch.org/fact-tank/2023/01/03/118th-congress-has-a-record-number-of-women/

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic Perspectives*, *21*(2), 153-174.

Lodge, M., & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology*, *26*(3), 455-482.

Mendelberg, T. (2001). *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton University Press.

Milgram, S., Mann, L., & Harter, S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly*, *29*(3), 437.

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science*, 23(7), 710-717.

Mo, C. H. (2015). The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Political Behavior*, *37*(2), 357-395.

Montgomery, N. V., & Cowen, A. P. (2020). How leader gender influences external audience response to organizational failures. Journal of Personality and Social Psychology, 118(4), 639-660.

Newman, B., Merolla, J. L., Shah, S., Lemi, D. C., Collingwood, L., & Ramakrishnan, S. K. (2021). The Trump effect: An experimental investigation of the emboldening effect of racially inflammatory elite communication. *British Journal of Political Science, 51*(3), 1138-1159.

Okimoto, T. G., & Brescoll, V. L. (2010). The price of power: Power seeking and backlash against female politicians. *Personality and Social Psychology Bulletin*, *36*(7), 923-936.

Ono, Y., & Burden, B. C. (2019). The contingent effects of candidate sex on voter choice. *Political Behavior*, *41*(3), 583-607.

Pager, D. (2003). The Mark of a Criminal Record. *American Journal of Sociology*, *108*(5), 937-975.

Piliavin, I. M., Rodin, J., & Piliavin, J. A. (1969). Good samaritanism: an underground phenomenon?. *Journal of Personality and Social Psychology*, *13*(4), 289.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811.

Presser, S., & Stinson, L. (1998). Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review*, 137-145.
Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870-10875.

Reid, L. D., & Foels, R. (2010). Cognitive complexity and the perception of subtle racism. Basic and Applied Social Psychology, 32(4), 291-301.

Saccardo, S. (2015). *Essays in Behavioral Economics and Ethics* (Doctoral dissertation, UC San Diego).

Schaffner, B. F. (2020). *The acceptance and expression of prejudice during the Trump era.* Cambridge University Press.

Schwarz, S., & Coppock, A. (2022). What have we learned about gender from candidate choice experiments? A meta-analysis of sixty-seven factorial survey experiments. *The Journal of Politics*, *84(2)*, 655-668.

Sevi, S., Arel-Bundock, V., & Blais, A. (2019). Do Women Get Fewer Votes? No. *Canadian Journal of Political Science/Revue Canadienne de Science Politique*, *52*(1), 201-210.

Simonsohn, U. (2016). Greg vs. Jamal: Why didn't Bertrand and mullainathan (2004) replicate. *Data Colada*, *51*.

Smith, J. L., Paul, D., & Paul, R. (2007). No place for a woman: Evidence for gender bias in evaluations of presidential candidates. *Basic and Applied Social Psychology*, *29(3)*, 225-233.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Duinn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology*

Bulletin, *24*(11), 1139-1152.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(5), 701-717.

Taber, C. S., Lodge, M., & Glathar, J. (2001). The motivated construction of political judgments. In J. H. Kuklinski (Ed.), *Citizens and politics: Perspectives from political psychology*. London: Cambridge University Press.

Teele, D. L., Kalla, J., & Rosenbluth, F. (2018). The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics. *American Political Science Review*, *112(3)*, 525-541.

Tjaden, J. D., Schwemmer, C., & Khadjavi, M. (2018). Ride with me—Ethnic discrimination, social markets, and the sharing economy. *European Sociological Review*, *34*(4), 418-432.

Trope, Y., & Liberman, N. (2003). Temporal Construal. *Psychological Review*, *110(3)*, 403-421.

Westwood, S. J., & Peterson, E. (2020). The inseparability of race and partisanship in the United States. *Political Behavior*, *44(3)*, 1-23.

Zitek, E. M., & Hebl, M. R. (2007). The role of social norm clarity in the influenced expression of prejudice over time. *Journal of Experimental Social Psychology*, *43*(6), 867-876.

A.4 Appendices

A.4.1 Randomisation checks for participants by condition

Here I consider the similarity of participant attributes across conditions, in order to test for successful randomisation of participants into experimental controls and treatments. Randomisation implies that participants should be similar in observed and unobserved characteristics between conditions, which facilitates causal inference from relationships between variables (Cox, 2009). If randomisation has occurred successfully, means demographic variables on which the sample is balanced should not to vary significantly between conditions. As can be seen, there are no significant differences amongst participants on key characteristics across conditions.

A.4.1 Randomisation checks for participants by condition

	Age	Gender	Region	Partisanship
(Intercept = White male, no scandal)	47.62 ***	1.56 ***	2.57 ***	4.90 ***
	(0.00)	(0.00)	(0.00)	(0.00)
White female, no scandal	-4.27	-0.05	0.08	-0.71
	(0.09)	(0.47)	(0.60)	(0.15)
White male, scandal	-1.01	-0.07	-0.05	-0.26
	(0.69)	(0.34)	(0.76)	(0.59)
White female, scandal	-0.63	-0.05	-0.01	-0.23
	(0.80)	(0.47)	(0.97)	(0.64)
Black male, no scandal	-1.15	-0.06	0.07	0.01
	(0.64)	(0.45)	(0.65)	(0.98)
Black female, no scandal	-2.36	0.01	-0.09	0.41
	(0.35)	(0.87)	(0.58)	(0.40)
Black male, scandal	-1.73	-0.04	-0.02	-0.06
	(0.49)	(0.56)	(0.90)	(0.91)
Black female, scandal	-2.84	-0.08	0.12	0.19
	(0.26)	(0.30)	(0.44)	(0.70)
N	709	709	709	709
R2	0.01	0.00	0.00	0.01

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. Partisanship is coded on a 1 to 10 scale where 1 = Strong Democrat and 10 = Strong Republican. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

A.4.2 Covariate analysis for DVs in controls

A.4.2 presents linear regression models for favourability and voting likelihood in the control conditions, including key covariates and an interaction between condition and partisanship.

	Favourability	Vote for
(Intercept = White male, no scandal)	6.65 ***	6.87 ***
	(0.00)	(0.00)
Condition = White female, no scandal	0.44	0.19
	(0.31)	(0.71)
Condition = Black male, no scandal	0.11	-0.11
	(0.79)	(0.83)
Condition = Black female, no scandal	0.54	0.71
	(0.26)	(0.20)
Age	0.00	-0.00
	(0.82)	(0.76)
Gender = Female	0.69 ***	0.65 **
	(0.00)	(0.00)
Party ID = Independent	-0.01	-0.48
	(0.98)	(0.34)
Party ID = Republican	-0.45	0.09
	(0.35)	(0.86)
White	0.18	
	(0.43)	
Region = Midwest	-0.13	-0.02
	(0.67)	(0.97)
Region = South	-0.01	-0.22
	(0.97)	(0.46)
Region = West	0.13	0.19
	(0.65)	(0.56)
White female, no scandal x Independent	-0.82	-0.30

	(0.19)	(0.67)
Black male, no scandal x Independent	-0.19	0.10
	(0.76)	(0.89)
Black female, no scandal x Independent	-0.48	-0.36
	(0.45)	(0.63)
White female, no scandal x Republican	-0.19	-0.43
	(0.79)	(0.59)
Black male, no scandal x Republican	0.70	-0.05
	(0.28)	(0.94)
Black female, no scandal x Republican	-0.85	-1.52
	(0.21)	(0.05)
N	362	362
R2	0.07	0.06

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

A.4.3 Covariate analysis for DVs in treatments

A.4.3 presents linear regression models for favourability and voting likelihood in the treatment conditions, including key covariates and an interaction between condition and partisanship.

	Favourability	Vote for
(Intercept = White male, scandal)	5.45 ***	5.46 ***
	(0.00)	(0.00)
Condition = White female, scandal	0.01	-0.20
	(0.99)	(0.76)
Condition = Black male, scandal	-0.57	-0.91
	(0.38)	(0.18)
Condition = Black female, no scandal	1.40 *	1.60 *
	(0.04)	(0.02)
Age	0.00	0.01
	(0.64)	(0.47)
Gender = Female	0.08	-0.09
	(0.78)	(0.77)
Party ID = Independent	-0.25	-0.70
	(0.71)	(0.33)
Party ID = Republican	0.84	0.50
	(0.23)	(0.49)
White	-0.43	-0.36
	(0.23)	(0.33)
Region = Midwest	-0.73	-0.87
	(0.10)	(0.06)
Region = South	-0.44	-0.48
	(0.28)	(0.25)
Region = West	-0.35	-0.75
	(0.41)	(0.09)

White female, scandal x Independent	0.41	1.29
	(0.67)	(0.21)
Black male, scandal x Independent	0.64	1.34
	(0.51)	(0.18)
Black female, scandal x Independent	-1.60	-1.93
	(0.11)	(0.07)
White female, scandal x Republican	-0.28	0.03
	(0.77)	(0.98)
Black male, scandal x Republican	-1.28	-0.70
	(0.20)	(0.49)
Black female, scandal x Republican	-3.00 **	-3.16 **
	(0.00)	(0.00)
N	347	347
R2	0.08	0.09

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West.
*** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.5 Tiered covariate models for Favourability, T1 vs T3, all races

A.4.5 presents the full results of tiered covariate models for favourability between T1 and T3, an abridged version of which are presented in the main text.

	Model 1	Model 2	Model 3	Model 4
(Intercept = White male, scandal)	5.32 ***	5.33 ***	4.02 *	3.39
	(0.00)	(0.00)	(0.03)	(0.08)
Condition = Black male, scandal	-0.83 *	-0.80	-0.95 *	-0.53
	(0.04)	(0.05)	(0.03)	(0.44)
Age	-0.01	-0.01	-0.01	-0.00
	(0.54)	(0.57)	(0.64)	(0.75)
Gender = female	0.50	0.48	0.67	0.69
	(0.22)	(0.25)	(0.12)	(0.10)
Region = Midwest		0.03	0.03	-0.01
		(0.96)	(0.96)	(0.99)
Region = South		-0.36	-0.29	-0.24
		(0.52)	(0.60)	(0.67)
Region = West		-0.03	0.14	-0.01
		(0.95)	(0.81)	(0.98)
Party ID = Independent		0.10	0.25	0.14
		(0.84)	(0.63)	(0.84)
Party ID = Republican		0.20	0.41	1.16
		(0.70)	(0.44)	(0.11)
Race = Asian or Pacific Islander			1.48	2.09
			(0.41)	(0.25)
Race = Black or African American			2.43	2.76
			(0.17)	(0.12)
Race = Hispanic or Latino			-0.62	-0.13
			(0.74)	(0.95)
Race = Multiracial or Biracial			-0.45	0.02

			(0.84)	(0.99)
Race = Native American			2.63	3.17
			(0.29)	(0.21)
Race = White or Caucasian			1.02	1.37
			(0.52)	(0.40)
Black male, scandal x Independent				0.14
				(0.89)
Black male, scandal x Republican				-1.54
				(0.14)
N	177	177	177	177
R2	0.03	0.04	0.09	0.10

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West.
*** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.6 Tiered covariate models for Favourability, T1 vs T3, race binary

A.4.6 presents the same models as above with the 'race' variable collapsed to a binary of white or non-white participants. Results indicate no significant differences between specifying the variable either way.

	Model 1	Model 2	Model 3	Model 4
(Intercept = White male, scandal)	5.32 ***	5.33 ***	5.41 ***	5.28 ***
	(0.00)	(0.00)	(0.00)	(0.00)
Condition = Black male, scandal	-0.83 *	-0.80	-0.80	-0.62
	(0.04)	(0.05)	(0.06)	(0.36)
Age	-0.01	-0.01	-0.01	-0.00
	(0.54)	(0.57)	(0.60)	(0.71)
Gender = female	0.50	0.48	0.48	0.52
	(0.22)	(0.25)	(0.26)	(0.22)
Region 2 = Midwest		0.03	0.04	-0.01
		(0.96)	(0.94)	(0.99)
Region 3 = South		-0.36	-0.34	-0.26

		(0.52)	(0.55)	(0.65)
Region 4 = West		-0.03	-0.05	-0.19
		(0.95)	(0.93)	(0.75)
Party ID = Independent		0.10	0.10	-0.19
		(0.84)	(0.84)	(0.79)
Party ID = Republican		0.20	0.22	0.84
		(0.70)	(0.67)	(0.25)
White			-0.16	-0.22
			(0.76)	(0.66)
Black male, scandal x Independent				0.57
				(0.57)
Black male, scandal x Republican				-1.26
				(0.22)
N	177	177	177	177
R2	0.03	0.04	0.04	0.05

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.7 Control by control comparisons, main effects

White male versus white female (Control, no covariates)

	Favourability	Vote for
(Intercept = White male, no scandal)	7.09 ***	6.98 ***
	(0.00)	(0.00)
Condition = White female, no scandal	0.05	-0.06
	(0.85)	(0.85)
N	178	178
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black male (Control, no covariates)

	Favourability	Vote for
(Intercept = White male, no scandal)	7.09 ***	6.98 ***
	(0.00)	(0.00)
Condition = Black male, no scandal	0.21	-0.07
	(0.45)	(0.82)
N	186	186
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black female (Control, no covariates)

	Favourability	Vote for
(Intercept = White male, no scandal)	7.09 ***	6.98 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	0.04	0.07
	(0.90)	(0.83)
N	180	180
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White female versus Black female (Control, no covariates)

	Favourability	Vote for
(Intercept = White female, no scandal)	7.14 ***	6.92 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	-0.01	0.13
	(0.96)	(0.67)
N	176	176
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

Black male versus Black female (Control, no covariates)

	Favourability	Vote for
(Intercept = Black male, no scandal)	7.29 ***	6.91 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	-0.17	0.14
	(0.52)	(0.64)
N	184	184
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.8 Control by control comparisons, with covariates and partisan interaction

White male versus white female (Control, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, no scandal)	5.88 ***	6.23 ***
	(0.00)	(0.00)
Condition = White female, no scandal	0.41	0.12
	(0.35)	(0.81)
Age	0.01	0.01
	(0.26)	(0.47)
Gender = female	0.58 *	0.47
	(0.04)	(0.16)
Region 2 = Midwest	0.30	0.65
	(0.50)	(0.22)
Region 3 = South	0.17	-0.14
	(0.67)	(0.76)
Region 4 = West	0.28	0.42
	(0.52)	(0.41)
Party ID = Independent	-0.09	-0.59
	(0.84)	(0.26)

Party ID = Republican	-0.68	-0.11
	(0.17)	(0.85)
White	0.61	0.33
	(0.07)	(0.42)
White female, no scandal x Independent	-0.75	-0.17
	(0.23)	(0.82)
White female, no scandal x Republican	-0.05	-0.33
	(0.95)	(0.69)
N	178	178
R2	0.07	0.06

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

White male versus Black male (Control, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, no scandal)	6.37 ***	6.61 ***
	(0.00)	(0.00)
Condition = Black male, no scandal	0.17	-0.11
	(0.71)	(0.83)
Age	0.01	0.01
	(0.27)	(0.47)
Gender = female	0.29	0.24
	(0.30)	(0.46)
Region 2 = Midwest	-0.46	-0.00
	(0.31)	(1.00)
Region 3 = South	0.04	-0.19
	(0.92)	(0.67)
Region 4 = West	-0.01	0.11
	(0.98)	(0.83)
Party ID = Independent	-0.04	-0.54
	(0.93)	(0.32)
Party ID = Republican	-0.55	-0.02
	(0.30)	(0.97)
White	0.40	0.18
	(0.25)	(0.66)
Black male, no scandal x Independent	-0.19	0.15
	(0.78)	(0.85)
Black male, no scandal x Republican	0.59	-0.07
	(0.39)	(0.93)
N	186	186
R2	0.05	0.02

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

White male versus Black female (Control, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, no scandal)	6.34 ***	6.73 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	0.54	0.72
	(0.28)	(0.21)
Age	0.00	-0.00
	(0.79)	(0.89)
Gender = female	0.72 *	0.47
	(0.01)	(0.15)
Region 2 = Midwest	0.03	0.17
	(0.95)	(0.74)
Region 3 = South	0.17	0.01
	(0.64)	(0.98)
Region 4 = West	0.28	0.40
	(0.52)	(0.42)
Party ID = Independent	-0.02	-0.52
	(0.97)	(0.33)
Party ID = Republican	-0.52	0.01
	(0.32)	(0.98)
White	0.33	0.16
	(0.39)	(0.71)
Black female, no scandal x Independent	-0.47	-0.34
	(0.49)	(0.66)
Black female, no scandal x Republican	-0.83	-1.53
	(0.25)	(0.06)
N	180	180
R2	0.08	0.06

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

White female versus Black female (Control, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White female, no scandal)	7.19 ***	7.48 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	0.01	0.43
	(0.98)	(0.40)
Age	-0.00	-0.01
	(0.77)	(0.30)
Gender = female	1.05 ***	1.06 ***
	(0.00)	(0.00)
Region 2 = Midwest	0.13	-0.17
	(0.74)	(0.71)
Region 3 = South	-0.04	-0.26
	(0.89)	(0.50)
Region 4 = West	0.27	0.25
	(0.48)	(0.57)
Party ID = Independent	-0.86 *	-0.84
	(0.04)	(0.07)
Party ID = Republican	-0.73	-0.39
	(0.12)	(0.47)
White	-0.08	-0.31
	(0.80)	(0.39)
Black female, no scandal x Independent	0.42	0.02
	(0.47)	(0.98)
Black female, no scandal x Republican	-0.45	-0.89
	(0.48)	(0.23)
N	176	176
R2	0.16	0.15

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

Black male versus Black female (Control, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = Black male, no scandal)	7.20 ***	7.42 ***
	(0.00)	(0.00)
Condition = Black female, no scandal	0.39	0.75
	(0.40)	(0.15)
Age	-0.00	-0.01
	(0.63)	(0.26)
Gender = female	0.83 **	0.83 **
	(0.00)	(0.00)
Region 2 = Midwest	-0.39	-0.53
	(0.34)	(0.25)
Region 3 = South	0.00	-0.12
	(1.00)	(0.76)
Region 4 = West	0.13	0.13
	(0.74)	(0.77)
Party ID = Independent	-0.20	-0.39
	(0.66)	(0.43)
Party ID = Republican	0.37	0.11
	(0.40)	(0.83)
White	-0.23	-0.40
	(0.48)	(0.28)
Black female, no scandal x Independent	-0.24	-0.43
	(0.71)	(0.55)
Black female, scandal x Republican	-1.48 *	-1.34
	(0.02)	(0.06)
N	184	184

R2	0.11	0.11

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.9 Treatment by treatment comparisons, main effects

White male versus white female (Treatment, no covariates)

	Favourability	Vote for
(Intercept = White male, scandal)	5.22 ***	4.93 ***
	(0.00)	(0.00)
Condition = White female, scandal	0.01	0.16
	(0.99)	(0.71)
N	173	173
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black male (Treatment, no covariates)

	Favourability	Vote for
(Intercept = White male, scandal)	5.22 ***	4.93 ***
	(0.00)	(0.00)
Condition = Black male, scandal	-0.81 *	-0.74
	(0.05)	(0.08)
N	177	177
R2	0.02	0.02

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black female (Treatment, no covariates)

	Favourability	Vote for
(Intercept = White male, scandal)	5.22 ***	4.93 ***

	(0.00)	(0.00)
Condition = Black female, scandal	-0.24	-0.21
	(0.57)	(0.62)
N	173	173
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

White female versus Black female (Treatment, no covariates)

	Favourability	Vote for
(Intercept = White female, scandal)	5.22 ***	5.09 ***
	(0.00)	(0.00)
Condition = Black female, scandal	-0.25	-0.38
	(0.55)	(0.38)
N	170	170
R2	0.00	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

Black male versus Black female (Treatment, no covariates)

	Favourability	Vote for
(Intercept = Black male, scandal)	4.40 ***	4.19 ***
	(0.00)	(0.00)
Condition = Black female, scandal	0.57	0.53
	(0.16)	(0.22)
N	174	174
R2	0.01	0.01

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.10 Treatment by treatment comparisons, with covariates and partisan interaction

White male versus white female (Treatment, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, scandal)	5.95 ***	6.53 ***
	(0.00)	(0.00)
Condition = White female, scandal	-0.00	-0.25
	(0.99)	(0.71)
Age	0.01	-0.00
	(0.69)	(0.81)
Gender = female	-0.74	-0.97 *
	(0.08)	(0.03)
Region 2 = Midwest	-0.29	-0.67
	(0.64)	(0.31)
Region 3 = South	-0.04	-0.14
	(0.94)	(0.81)
Region 4 = West	-0.04	-0.72
	(0.94)	(0.24)
Party ID = Independent	-0.28	-0.83
	(0.69)	(0.26)
Party ID = Republican	0.91	0.51
	(0.20)	(0.49)
White	-1.00 *	-0.71
	(0.04)	(0.17)
White female, scandal x Independent	0.40	1.38
	(0.69)	(0.19)
White female, scandal x Republican	-0.33	-0.01
	(0.74)	(0.99)
N	173	173
R2	0.06	0.07

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black male (Treatment, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, scandal)	5.28 ***	5.29 ***
	(0.00)	(0.00)
Condition = Black male, scandal	-0.62	-1.00
	(0.36)	(0.16)
Age	-0.00	-0.00
	(0.71)	(0.80)
Gender = female	0.52	0.25
	(0.22)	(0.57)
Region 2 = Midwest	-0.01	-0.01
	(0.99)	(0.98)
Region 3 = South	-0.26	-0.44
	(0.65)	(0.46)
Region 4 = West	-0.19	-0.75
	(0.75)	(0.22)
Party ID = Independent	-0.19	-0.66
	(0.79)	(0.38)
Party ID = Republican	0.84	0.41
	(0.25)	(0.59)
White	-0.22	0.13
	(0.66)	(0.81)
Black male, scandal x Independent	0.57	1.30
	(0.57)	(0.22)
Black male, scandal x Republican	-1.26	-0.63
	(0.22)	(0.55)
N	177	177
R2	0.05	0.05

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

White male versus Black female (Treatment, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White male, scandal)	6.79 ***	6.47 ***
	(0.00)	(0.00)
Condition = Black female, scandal	1.29	1.42 *
	(0.06)	(0.04)
Age	-0.02	-0.01
	(0.21)	(0.37)
Gender = female	0.20	-0.00
	(0.63)	(1.00)
Region 2 = Midwest	-1.48 *	-1.10
	(0.02)	(0.09)
Region 3 = South	-0.50	0.12
	(0.40)	(0.84)
Region 4 = West	-0.42	-0.76
	(0.48)	(0.22)
Party ID = Independent	-0.34	-0.85
	(0.63)	(0.23)
Party ID = Republican	0.95	0.56
	(0.18)	(0.43)
White	-0.73	-0.65
	(0.14)	(0.20)
Black female, scandal x Independent	-1.29	-1.57
	(0.21)	(0.13)
Black female, scandal x Republican	-2.89 **	-3.24 **
	(0.00)	(0.00)
N	173	173
R2	0.14	0.16

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

White female versus Black female (Treatment, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = White female, scandal)	5.48 ***	5.21 ***
	(0.00)	(0.00)
Condition = Black female, scandal	1.38 *	1.75 **
	(0.04)	(0.01)
Age	0.02	0.03
	(0.15)	(0.06)
Gender = female	-0.37	-0.45
	(0.35)	(0.27)
Region 2 = Midwest	-1.60 *	-1.82 **
	(0.01)	(0.01)
Region 3 = South	-0.76	-0.74
	(0.19)	(0.21)
Region 4 = West	-0.55	-0.71
	(0.39)	(0.28)
Party ID = Independent	0.19	0.60
	(0.78)	(0.39)
Party ID = Republican	0.62	0.72
	(0.37)	(0.30)
White	-0.66	-0.92
	(0.19)	(0.07)
Black female, scandal x Independent	-1.81	-2.87 **
	(0.07)	(0.00)
Black female, scandal x Republican	-2.69 **	-3.13 **
	(0.01)	(0.00)
N	170	170
R2	0.14	0.20

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

Black male versus Black female (Treatment, with covariates & Condition x Party interaction)

	Favourability	Vote for
(Intercept = Black male, scandal)	4.54 ***	3.67 ***
	(0.00)	(0.00)
Condition = Black female, scandal	2.19 **	2.62 ***
	(0.00)	(0.00)
Age	-0.00	0.01
	(0.81)	(0.34)
Gender = female	0.90 *	0.78
	(0.02)	(0.06)
Region 2 = Midwest	-1.54 *	-1.28
	(0.02)	(0.05)
Region 3 = South	-0.91	-0.76
	(0.11)	(0.21)
Region 4 = West	-0.80	-0.71
	(0.20)	(0.27)
Party ID = Independent	0.53	0.78
	(0.43)	(0.26)
Party ID = Republican	-0.50	-0.26
	(0.46)	(0.72)
White	0.39	0.02
	(0.45)	(0.96)
Black female, scandal x Independent	-2.44 *	-3.40 **
	(0.01)	(0.00)
Black female, scandal x Republican	-1.80	-2.39 *
	(0.07)	(0.02)
N	174	174

R2	0.15	0.16

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. White is coded as 0 = non-white, 1 = white. *** p < 0.001; ** p < 0.01; * p < 0.05.

A.4.11 Additional results of observational analyses

Difference between vote share in previous and subsequent election by scandal type

	Model 1
(Intercept)	0.10 ***
	(0.00)
Abuse of office	0.03
	(0.54)
Bribery	0.08
	(0.07)
Campaign finance	-0.10
	(0.18)
Campaign violations	-0.07
	(0.39)
Inadequate disclosure	-0.02
	(0.78)
Lying	-0.03
	(0.82)
Sex	-0.01
	(0.82)
Sexism/Racism	-0.00
	(0.99)
Spying	-0.11
	(0.34)
Unsavory behavior	-0.10
	(0.26)
Violence	-0.15
	(0.08)

N	65
R2	0.23

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

Expanded version of Table 4.2 in main text

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	0.07 ***	-0.08	-0.10	-0.11	-0.07
	(0.00)	(0.38)	(0.29)	(0.25)	(0.57)
Minority = 1	0.07 *	0.08 *	0.08 **	0.08 *	0.06
	(0.02)	(0.01)	(0.01)	(0.02)	(0.11)
Seniority		-0.00	-0.00	-0.00	-0.01
		(0.57)	(0.70)	(0.43)	(0.35)
LES		0.00	-0.00	-0.00	-0.00
		(0.96)	(0.84)	(0.83)	(0.97)
Age		0.00	0.00	0.00	0.00
		(0.14)	(0.16)	(0.09)	(0.26)
Media coverage			0.03	0.04	0.05
			(0.32)	(0.26)	(0.21)
Senator				-0.05	-0.03
				(0.17)	(0.38)
Abuse of office					0.02
					(0.73)
Bribery					0.07
					(0.13)
Campaign Finance					-0.10
					(0.19)
Campaign violations					-0.03
					(0.73)

Inadequate disclosure					-0.01
					(0.91)
Lying					-0.00
					(0.99)
Sex					-0.04
					(0.52)
Sexism / Racism					-0.00
					(0.99)
Spying					-0.09
					(0.43)
Unsavoury behaviour					-0.09
					(0.27)
Violence					-0.15
					(0.10)
N	65	65	65	65	65
R2	0.09	0.13	0.14	0.17	0.33

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

A.4.12 Examples of control and treatment survey instruments

Survey instrument C1

MV Local Zachary Miller Announces Candidacy for State House

MOUNT VERNON, OH — June 21, 2021 — *The Daily Herald* can confirm that Columbus native Zachary Miller has announced his candidacy for the state House of Representatives in Ohio's 19th district. Miller is a graduate of The Ohio State University and OSU's College of Law, where he received his Juris Doctorate (JD) and was associate editor of the Legal Review.

Mr. Miller has been a partner at the Columbus-based law firm of Stevens & Seigel since 2011, and has twice won the Knox County Law Association's Distinguished Service Award for legal service to the local community for pro-bono legal work. Asked about his decision to leave his legal career and run for office, the Miller campaign replied:

"Zach has always felt passionately about serving his community. Inspired by the examples of hard work and public service set by his father, a firefighter, and mother, an emergency room nurse, he made the difficult decision to set aside the career he loves and try and fix the polarization and inaction of Ohio's legislature."

Mr. Miller, 38, is also a senior non-resident fellow at the public policy advocacy group OhioTech, where, as a citizen, he has worked with local, state, and federal lawmakers to craft legislation on telecommunications, broadband deployment, energy and energy efficiency, grid modernization, education, job creation, environmental protection, food safety, seismic safety, and a host of other issues. He has stated that his campaign will be focused on securing improved funding for education, job training, and health care for the people of Knox County.

Mr. Miller lives in Mount Vernon with his wife Hannah and their three children, and is an avid Buckeyes fan.

Survey instrument T2

MV Local Hannah Miller Announces Candidacy for State House

MOUNT VERNON, OH — June 21, 2020 — *The Daily Herald* can confirm that Columbus native Hannah Miller has announced her candidacy for the state House of Representatives in Ohio's 19th district. Miller is a graduate of The Ohio State University and OSU's College of Law, where she received her Juris Doctorate (JD) and was associate editor of the Legal Review.

Ms. Miller has been a partner at the Columbus-based law firm of Stevens & Seigel since 2011, and has twice won the Knox County Law Association's Distinguished Service Award for legal service to the local community for pro-bono legal work. Asked about her decision to leave her legal career and run for office, the Miller campaign replied:

"Hannah has always felt passionately about serving her community. Inspired by the examples of hard work and public service set by her father, a firefighter, and mother, an emergency room nurse, made the difficult decision to set aside the career she loves and try and fix the polarization and inaction of Ohio's legislature."

Ms. Miller, 38, is also a senior non-resident fellow at the public policy advocacy group OhioTech, where, as a citizen, she has worked with local, state, and federal lawmakers to craft legislation on telecommunications, broadband deployment, energy and energy efficiency, grid modernization, education, job creation, environmental protection, food safety, seismic safety, and a host of other issues. She has stated that her campaign will be focused on securing improved funding for education, job training, and health care for the people of Knox County.

Ms. Miller lives in Mount Vernon with her husband Zach and their three children, and is an avid Buckeyes fan.

[...]

State House Hopeful Hannah Miller Investigated for Owed Taxes

MOUNT VERNON, OH — July 2, 2021 — *The Knox Tribune* can today confirm that a recent political ad campaign against the candidacy of Ms. Hannah Miller (OH-19) is factually accurate.

The state of Ohio has been forced to file 9 "tax warrants" against Ms. Miller between 2004 and 2016, totaling an outstanding tax debt of some $19,251.09. A call to Ohio tax authorities indicates that with penalties and interest, Ms. Miller's unpaid bill amounts to $27,634.29.

The Daily Herald reported yesterday that Ms. Miller learned about the warrant from the recent advertisement and paid it on Wednesday, June 17, 2021. By contrast, on Thursday, June 18, 2021, Ms. Miller told the *Herald* that she paid her outstanding tax debts in January at the time she learned about them.

Speaking with local officials, the *Tribune* understands that this pattern of tax delinquency could rise to the level of a misdemeanor offense, though it is unclear at this time whether any charges are being considered against Ms. Miller.

The *Tribune* has reached out to the Miller campaign for comment, and will update this article accordingly if and when a response is received.

5

5 Give US the Ballot? Dynamism in Beliefs on Voting Access

This paper is focused on Americans' beliefs about the role of the state in providing voting access to its citizens. It addresses a potential information asymmetry resulting from divergent voting experiences: while most voters in the US wait less than ten minutes to cast their ballots, a small but meaningful minority (roughly 5 million citizens) face queues of more than an hour. This *time tax* (Mukherjee, 2009) is a significant deterrent to present and future voting (e.g., leaving the queue without voting or not voting in a subsequent election) — and lost votes can bear significantly on final results. Moreover, long wait times are not evenly distributed, such that minority voters, voters in more densely populated areas, and poorer voters wait longer. I explore beliefs about the acceptability of such impediments to voting, support for policy reform aimed at addressing them, and, crucially, how dynamic these beliefs are to information and other contextual factors. I address these questions with data from a novel survey experiment (n = 988) with five conditions. The first is a baseline survey of Americans' views on wait times to vote and potential reform, and serves as the control for subsequent conditions. It is the first of its kind to separate beliefs about voting ease (i.e., how long one waits to cast their ballot) from issues of eligibility or registration (e.g., voter ID laws), focusing only on impediments at the point of voting, an area that has received considerably less attention than other voting restrictions (Pettigrew, 2021). Further experimental conditions consider how responsive people's beliefs are to information about long queues, and if and how partisan or race-based prejudices affect propensity to support reform. I find the majority of respondents across the political spectrum generally favour reforms to curb long wait times — and that this policy preference is dynamic to new information such that the treatment significantly increases that support, though the effect varies slightly based on who is depicted facing long waits.

5.1 Introduction

Long wait times to vote are a recurrent feature of US elections, but only recently has academic research focused on systematically analysing the causes and consequences of such waits (Stewart, 2014). Those efforts have yielded both positive and negative findings: on the one hand, the average American queues for less than 10 minutes to cast their ballot in a typical election, and roughly 95% of voters in the 2014 midterm elections described their experience at the polls as "excellent" or "very good" (Stewart, 2015). Conversely, in the 2012 presidential election, more than five million voters — or roughly 4% of the entire electorate — waited more than an hour to vote (Presidential Commission on Election Administration, 2014), and wait times in states like Georgia regularly exceed this threshold, sometimes taking voters as long as 10 hours to reach the voting booth (BBC News, 2020).

Research suggests these wait times have a tangible effect on voters' behaviour in both present and future elections. For example, MIT's Survey on the Performance of American Elections estimates that long waits alone deterred three million registered voters from casting their ballots in the 2014 midterms, with people either abandoning the queue or not turning up altogether because of the time it would take to vote (Stewart, 2015). In total, they calculate some 15 million votes were *lost* in that election cycle due to issues related to election administration, rather than individuals' motivation to participate. As for the downstream consequences, Pettigrew (2021) finds that every one-hour increase in wait times at one's voting precinct decreases voters' propensity to vote in subsequent elections by a percentage point.

That wait times bear on actual propensity to vote — the central function of a democratic political system — threatens a hypothesised link between representative democracy and political accountability (i.e., democratic accountability). Barriers that affect turnout (and, conceivably, electoral outcomes) further strain the already "relatively weak" (Warren, 2014, p.7) ability of citizens as principles in a democratic system to hold politicians (their agents) accountable by voting. Yet, while a large literature on voting behaviour and access is devoted to other election-

related issues (e.g., voter ID laws), to date little is known about how Americans feel about impediments to voting at the point of voting.

This paper fills that gap by providing a novel survey of Americans' beliefs about voting wait times, and experimentally analysing how dynamic these beliefs are to information and other contextual factors such as the identity of disenfranchised voters. The results of this study serve two purposes: first, the baseline survey serves as a standalone resource to academics and policymakers interested in understanding how the general population regards issues relating to voting access; second, the experimental treatments draws on research in political psychology about when, and when not, people's political and policy preferences may be open to change, and empirically tests hypotheses generated from past research to better understand how and in what directions beliefs about voting access can (or cannot) change based on new information.

To summarise the main findings, first, Americans across the political spectrum generally support election reforms aimed at curbing long wait times to vote. Second, the information treatment highlighting these waits significantly affects these beliefs, such that: information increases both support for reform and the perception that long waits are bad for democracy. Third, there are some differences between information treatments based on the race of voters depicted waiting — for example, the photo treatment highlighting only Black voters somewhat counteracts the positive effect of information on support for reform — but these are not consistent across outcome measures. Apart from these primary outcome measures, participants are further asked about their support for three popular policies aimed at making it easier for people to vote: making it easier to vote by mail, making election day a federal holiday, and lengthening the period of early voting. Whereas beliefs about impediments at the point of voting may be relatively unformed due to the aforementioned asymmetry in voting experiences, mail voting has become the subject of considerable public debate and partisan polarisation, and as such may be less subject to change. This is supported by the experimental data, which shows significant partisan differences in support for making it easier to vote by mail, and that these

differences are stable to condition effects. This suggests that part of the reason for dynamism in beliefs about voting wait times may be due to the relatively unpolarised nature of past discourse surrounding this issue.

The rest of the paper proceeds as follows. Part 3.3 is a literature review bringing together past research on election administration and voting-related beliefs to explain how impediments at the point of voting affect political participation, why it matters, and why these beliefs might be dynamic to information. Part 3.4 details the design and methodology of the experiment, while Part 3.5 presents the results thereof. Part 3.6 discusses the main findings, their contributions, and considers avenues for future research.

5.2 Literature Review

5.2.1 The State of the Wait

Following the 2012 election, in which images depicting hours-long queues outside voting locations spurred a frenzy of attention amongst policymakers, then-President Obama launched the bipartisan Presidential Commission on Election Administration, aimed in part at providing insights into the causes and consequences of these long waits, and potential remedies for future elections. The Commission brought together experts from across the political divide to holistically review how American elections function (and how this functioning affects the behaviour of voters), and to make concrete recommendations for policy and administrative reforms that could improve the quality of future elections.

Though the 2012 election may have brought the issue to a head, protracted (and unevenly distributed) wait times have long been a feature of American elections. For example, an analysis of data collected by the Pew Research Center found that, averaging across national elections in 2004, 2006, and 2008, roughly 13% of voters who cast their vote on Election Day waited between one and two hours to vote, and an additional 7% waited more than two hours (Mukherjee, 2009). In the 2004 election in particular, voters in the Ohio cities of Columbus, Cincinnati, and Cleveland faced waits of more than 10 hours, with a number abandoning the queues without casting their ballots (Powell & Slevin, 2004). Based on the margins of the election in that state, systems researchers Belenky and Larson (2006) write that these queues "might have decided or at least substantially affected" the outcome of the election (p.30).

It was against this backdrop that the Presidential Commission included wait times as a key facet of ensuring voting access to American citizens. Amongst its recommendations for improving American elections, the Commission "concluded that, as a general rule, no voter should have to wait more than half an hour in order to have an opportunity to vote" (PCEA, 2014, p.14). Noting that there is no single cause nor single solution to long lines, the report

nonetheless asserted that meeting this recommendation is possible and the problem of long lines is "solvable."

Whilst the report presents hope for substantial electoral improvements in the future, unfortunately these solutions have yet to take hold. In the 2018 midterm elections, for example, roughly 5% of all precincts across the US had average wait times of longer than 30 minutes (Fortier et al, 2018), while 18% of all voters faced such waits in the previous 2016 presidential election (Chen et al, 2022), far exceeding the Commission's recommendation. Similarly, early retrospectives of the 2020 election show that long lines — further exacerbated by pandemic-related constraints and historic turnout — have not abated. At present, then, wait times to vote are still a prominent feature of American elections.

5.2.2 Who Has to Wait?

While the above review makes clear that *some* American voters continue to face long waits, the experience of voting in the US varies significantly depending on one's location and identity. These systemic differences are part and parcel to the problem of voting access. Geographically, wait times vary dramatically both between and within states and counties across the country. In the 2018 election, voters in the best performing state of Vermont waited on average less than a minute, compared with voters in Georgia (that year's worst performing state), who faced an average wait time of 18 minutes (and counted a large share of all of the nation's hour-plus waits). Residents in the most densely populated neighbourhoods of the US waited 25% longer than those in the least populated areas (Fortier et al, 2018), disproportionately affecting urban voters who are in turn more likely to lean left, even after controlling for related socio-political characteristics such as race and economic status (Scala & Johnson, 2017).

Yet the greatest divides in ease of voting relate to demographic characteristics. In particular, voters' race and socioeconomic status (SES) are some of the most consistent predictors of their likely wait time. For example, in the 2008 election 27% of all Black voters reported facing long lines at their polling locations, compared to only 11% of white voters

(Alvarez et al, 2009). Voters in predominantly minority precincts face lines that are, on average, twice as long as those in predominantly white precincts, and minority voters are three and six times as likely to face waits of longer than 30 or 60 minutes, respectively, than are their white counterparts (approximately 6.7% of all Black voters wait more than an hour) (Pettigrew, 2017).

Splitting precincts into deciles based on their percentage of Black voters, Chen and co-authors (2022) find that voters located in areas in the top decile wait 19% longer than those in the bottom decile, and are 49% more likely to spend more than 30 minutes queuing. When the analysis compares entirely-Black to entirely-white precincts, the result is a 29% increase in average wait time and a 74% increased likelihood of waiting more than half an hour (Chen et al, 2022). Using a similar design, Fortier and co-authors (2018) find that the average wait time jumps from 5.1 minutes in precincts that are comprised of 90% or more white voters to 32.4 minutes in those that are 90% or more non-white. Summarising the persistence of such findings, Stein and co-authors write that race "has the biggest influence on waiting to vote" (Stein et al, 2020, p.431).

Beyond racial divides, individuals' SES also materially affects their experiences voting. For example, a *New York Times* investigation used location tracking data to analyse how neighbourhood income affected wait times in the 2020 election, concluding that voters in poorer neighbourhoods were more likely to face long wait times to vote than those in rich neighbourhoods (Quealy & Parlapiano, 2021). In 2018, voters in precincts with mean income of less than $40,000 waited on average 15.4 minutes to vote, compared with an only 7.7 minute average in other precincts (Fortier et al, 2018). Using a related measure of SES, homeownership, the authors find an average wait of 16.6 minutes in precincts with rental rates exceeding 60%, compared with 7.9 minutes in predominantly non-rental precincts (Fortier et al, 2018).

Taken together, it is clear that the experience of voting in American elections is rather different depending on one's place in society. That is, voting access — at least if construed of to include lost-time as a factor — is not uniformly distributed across the population. Rather, it is considerably easier for some Americans to vote than others, and these differences largely fall

along geographic, racial, and socioeconomic divides. As will be seen later, our experimental data similarly show divides in respondents' past voting experiences on the basis of these characteristics.

5.2.3 Do Wait Times Matter?

While the evidence reviewed thus far makes clear that long lines to vote are a persistent and unequally encountered feature of American elections, the obvious question that arises is: do these wait times affect democratic outcomes? This question is by no means settled given the nascent state of the literature, but there is a growing body of evidence that indicates wait times affect both individual voting behaviour — in the present, should voters choose not to wait to cast their ballots, and in subsequent elections, should this experience alter their decision to vote or not in the following election cycle — and in turn, these behaviour changes affect election results.

As for changing behaviour in the present, election post-mortems regularly find that long waits deter large numbers of citizens from ultimately casting their ballots. For example, it is estimated that at least a million eligible and registered voters chose not to vote in the 2000 election due to long lines at their precinct (Caltech / MIT Voting Technology Project, 2001). In 2014, this number tripled to roughly 3 million lost votes due to long lines (Stewart, 2015). This tracks with field evidence from California, where poll watchers found a 1.98% rate of reneging (i.e., leaving the queue) amongst all arrivals to voting locations (Spencer & Markovits, 2010).

While, of course, it is uncertain how those votes would have been cast, these lost votes often meet or exceed the determining margin of a given election, as was the case in Ohio in 2004, where an estimated 129,500 voters left due to long lines and the election was decided by a mere 118,000 votes (Mukherjee, 2009). Given the previously-reviewed evidence on systematic differences in wait times based on identity — and the relationships between identity and vote choice in the US (see, for example, Westwood & Peterson, 2020) — it is not unreasonable to

assume that, at the very least, lost votes would not have been evenly distributed between candidates and parties.

Moreover, these negative experiences waiting have downstream consequences: analysing vote history records, Pettigrew (2021) finds that for every one-hour increase in a voter's wait time, their propensity to vote in a subsequent election decreases by a percentage point. There are a few ways of interpreting why this may be the case. For one, as Pettigrew writes, wait times affecting future participation seems to track with theories of voting behaviour that view the decision to vote or not as the product of a cost-benefit analysis (see, for example, Riker and Ordeshook, 1968). Indeed, a broader literature correlates convenience and experiential elements of voting with subsequent behaviours. For example Gimpel and Schuknecht (2003) find that the ease and accessibility of polling locations significantly affects turnout, and Gimpel and co-authors (2006) find that proximity to early voting locations increases propensity to vote early. Put simply, from this perspective, the higher the cost is to casting one's ballot — in this case in the form of lost-time spent waiting — the less likely one will be to do so.

A second explanation leans more heavily on insights into voter psychology, and posits a relationship between voting experiences and perceptions of the system more broadly. For example, past research has found that administrative performance at the precinct-level affects voters perceptions of the fairness of elections (Bowler et al, 2015), and in turn, confidence in the fairness of the electoral process affects future voting propensity (Alvarez, Hall & Llewellyn, 2008). Similarly, Atkeson and Saunders (2007) show that the quality of an individual's experience voting affects their confidence in the voting process. Summarising why this might be the case, de Kadt (2017) posits that voting behaviour is in part driven by associations between past voting experiences and positive emotional states.

Whatever the explanation, the weight of academic evidence suggests that wait times to vote have a tangible effect on democratic participation in the US. In turn, this effect on participation — unevenly distributed amongst groups with higher likelihoods of facing long waits

— has the potential to change election outcomes, upon which the foundation of democratic accountability rests.

5.2.4 Americans' Beliefs on Voting Access

Even as wait times have been somewhat overlooked by the academic literature, a wide range of research documents Americans' views on other impediments to voting and proposed reforms, providing insight into the factors that most affect beliefs. Among the most common strains in this literature is research on voter ID laws — a much-discussed and partisan-polarised issue of voter preparedness. For example, Wilson and Brewer (2013) find that individuals' political views, as well as their racial attitudes, strongly influence their support of voter ID restrictions. Valentino and Neuner (2017) find that Democrats' dislike for such laws is so strong that enaction of voter ID restrictions mobilises Democratic turnout in subsequent elections.

Such a reaction may be in part due to the perception that these laws are undertaken with strategic interests in mind: research suggests that Republicans in states with lower levels of partisan competition are less likely to promote voter ID laws (Hicks et al, 2015). And this competitive jockeying is not one-sided: as Kane writes "Democrats' views toward VID are more rooted in strategic concerns about electoral outcomes than are Republicans" (2017, p.943) — and that, when told VID would reduce Republican turnout, partisan differences in support for VID disappeared entirely. As for the general public, Alvarez and co-authors (2011) find that Americans overwhelmingly support requirements to show photo ID before voting.

Other voting-related beliefs tend to split similarly along party lines. For example, a Pew Research Center survey from July of 2021 finds that, while Democrats tend to view voting as a fundamental right, Republicans are more likely to view it as a privilege that comes with responsibilities (Gomez & Doherty, 2021). During the 2020 election, as physical constraints related to the COVID-19 pandemic led to an uptick in voting by mail, a great deal of media and public attention was paid to the potential of fraud resulting from this form of voting. In

particular, Republican leaders in states like Pennsylvania, taking cues from President Trump, sought to invalidate allowances they had previously made for mail voting. During this time, Lockhart and co-authors (2021) observed a sharp widening in the partisan divide on views of the issue. This tracks with research showing that, as issues become more polarised, individuals become more fixed in their beliefs (Druckman, Peterson & Slothuus, 2013). Moreover, by finding a frame on the issue of mail voting that better resonated with their voters (from *making voting more convenient* to *preventing fraud*), Republican elites might have been more able to affect their co-partisan's views (Slothuus, 2010). However, by making the issue more salient, it is possible individuals would be less likely to lean solely on party cues (Ciuk & Yost, 2016).

As for the effects of race on beliefs, Wilson and co-authors (2014) show that white respondents who are exposed to photos of Black poll workers were more likely to support voter ID laws (the basis for one of the experimental treatments that follows). Similarly, Banks and Hicks (2015) find that inducing fear in respondents high in implicit racism led to greater support for voter ID restrictions, which they attribute to the activation of those attitudes. These effects of racial coding have been found in a number of political and policy contexts, for example on support for more punitive approaches to crime (Gilliam & Iyengar, 2000).

Taken together, this research shows that partisanship, identity, and competitive interests are all strong mediators of beliefs about voting in the US. As such, one might expect results to be somewhat polarised along party lines, and to find differences in the treatments based on identity and competitive interests. Because the issue of wait times is not already the subject of much partisan debate, it also seems reasonable to expect respondents of all political persuasions to be responsive to information.

5.2.5 Receptiveness of Beliefs to Information

Given the sharp disparity in wait times detailed above — with the modal voter experiencing efficient election day administration, and a small minority bearing the burden of long waits — it

seems likely there would be at least a slight asymmetry in people's awareness of the issue. This experiential gap means that most Americans may not have much considered the effects of long wait times on democratic participation and outcomes; and as such, that beliefs about wait times may be more amenable to changing based on new information than are other policy-related beliefs that are grounded in more salient or polarised issues.

But when and why are political beliefs more subject to change? A vast body of past research explores how people change (or do not change) their political beliefs in line with new information. One prominent theory of so-called *hot-cognition* contends that, once information about a given political actor or issue has been internalised, individuals' subsequent processing of information related to that issue will be primed with an automatic, affective association that in turn alters their consideration of new evidence (Lodge & Taber, 2000; Lodge & Taber, 2005). According to this line of thought, when political or policy issues are the subject of considerable past debate, people will be more likely to engage in motivated review of any subsequent information pertaining thereto, placing greater weight instead on these automatic cues to make their judgments (Burdein, Lodge & Taber, 2006). This accords with some broader theories of motivated information processing, which contend that strong prior beliefs are the most resistant to changing based on new evidence, whereas in cases where prior knowledge is limited, people may be more likely to assign higher judgment weights to new information (see, for example, Kunda, 1990; Kahneman & Tversky, 1983), or to process that information more thoroughly (Petty & Cacioppo, 1986).

In particular, research has found that the more prominent a political or policy issue is — and the more media coverage it receives (Jerit & Barabas, 2012) — the more likely individuals are to rely on broader partisan attitudes or prior beliefs rather than new information (Gaines et al, 2007). However, this is hardly a settled matter, as other scholars argue that more salient or important issues are more likely to prompt individuals to form their own views (Ciuk & Yost, 2016). Ultimately the extent to which people rely on prior beliefs versus holistically consider new

information likely depend on the individual and issue in question, and in particular rely on whether people see a *need* to evaluate new information. (Druckman & Nelson, 2003). Nonetheless, it seems clear that — as issues become more polarised along partisan dimensions — individuals' propensity tend to align their individual beliefs with the overall party cue (Levendusky, 2010; Mullinix, 2016).

As touched on above, this latter finding has recently been observed in the context of beliefs surrounding mail voting. As voting by mail became a topic of public scrutiny amidst the COVID-19 pandemic — with strong party-elite cues from Republican leaders like then-President Trump — Americans' beliefs became polarised along partisan lines. Whereas prior research had found cross-party similarity in opinions on voting by mail, a pair of representative surveys conducted ahead of the 2020 election found a 10 percentage point gap in opinions between Republicans and Democrats following a Trump tweet in April, which doubled to 20 percentage points by June of that year (Lockhart et al, 2020). This provides a nice contrast to the relatively undiscussed issue of wait times to vote, for which there is relatively less scholarship or public attention focusing on impediments at the point of voting (Pettigrew, 2021). Not that media stories covering long waits are uncommon, but, as research on the consequences of long waits has only recently begun to quantify the tangible effects they may have on actual voting behaviour and election outcomes, these stories may fail to provide the full context of the issue. Moreover, whereas there may be principled cases to make in favour of VID or against mail voting, it may be harder to construct a compelling justification in favour of long lines. Thus it seems likely that, relative to other voting beliefs, people's views on wait times may be more receptive to information, and that this may lead to relevant changes in policy preferences.

5.2.6 Research Puzzle

The present research attempts to address the information asymmetries that arise from Americans' vastly different voting experiences. Given the fact that the wide majority of voters

will not have faced long queues at polling stations, and therefore may not be aware of the burden they impose on democratic participation, the primary analysis focuses on whether and how the provision of information affects respondents' views on voting reform. As elucidated above, because wait times to vote are not the subject of prior polarisation, it is a context in which beliefs may be amenable to new information. Thus, the first hypothesis posits that:

H1: information treatment increases support for reform to address wait times

Of course, individuals' views on voting-related issues do not exist in a vacuum. As the research on VID laws reviewed above demonstrates, both racial prejudice and partisan competition seem to bear heavily on support for or opposition to reform. As such, the subsequent experimental treatments, reviewed further in 5.3.2, seek to evaluate whether and how these preferences alter the efficacy of the information treatment, hypothesising that:

H2: the identities of inconvenienced voters affects support for reform

H3: competitive stake in election outcomes affects support for reform

This context of wait times presents an ideal setting in which people's beliefs may be dynamic to changing based on the information presented. Put simply, lacking extensive previous consideration, polarisation, or a compelling issue frame, policy preferences may be less likely to be fixed. Moreover, because the burden of long waits fall so unevenly on minority and low-SES voters, the subsequent analyses provide crucial context as to whether and how the identities of inconvenienced voters (including the potential that those voters are out-partisans) may affect support for reform.

5.3 Experimental Design & Methods

5.3.1 Design

In order to interrogate people's beliefs about wait times, I design and administer a multi-treatment survey experiment (*n*=988). The survey is designed in Qualtrics and has five experimental conditions, described below. Participants are recruited through the sample provider Lucid Theorem, which provides US-based samples balanced for demographic and political characteristics. Previous research replicating experimental findings between Lucid and other online survey providers such as M-Turk finds that Lucid samples "track well with US national benchmarks" and that the platform is "suitable for evaluating many social scientific theories" (Coppock & McClellan, 2019). Once recruited, participants are redirected to the Qualtrics-based survey, where they must pass our conditions of informed consent before entering the survey. Immediately following the consent form, participants are randomised into an experimental condition and directed to a landing page with brief instructions for the survey to follow. The experimental conditions are as follows.

5.3.2 Experimental Conditions

C1: Baseline Survey

The first survey, which serves as the control condition for subsequent experiments, is meant to provide a baseline understanding of the public's opinions about wait times to vote, as well as a host of other election-related issues. Therein, participants answer the same series of questions about wait times that serve as the outcome variables of interest in the subsequent experimental conditions. The results of C1 as a standalone survey instrument provide a valuable resource for academics, election officials, and policymakers seeking to understand public opinion on impediments to voting *at the point of voting*. It is, to my knowledge, the first instrument of its type to gauge how the public feels about long lines to vote, bringing research in this area in line with that on other voting-related restrictions. As such, it can provide context to, for example, the

Presidential Commission's threshold of 30 minutes as the cut-off for voting wait times: is this recommendation in line with the public's view of what constitutes a reasonable wait?

As a control condition, the results of C1 also provide a benchmark against which subsequent experimental treatments are evaluated. In particular, I am interested in comparing outcomes on our main dependent variables (DVs, detailed further below) in C1 with those in the primary experimental condition, the information treatment (T1, detailed below). Results from C1 are further compared with the rest of the experimental treatments to determine whether they affect the relationship between information and beliefs.

T1: Information Treatment

The first experimental condition seeks to determine whether opinions about voting wait times are affected by knowledge thereof. Because the modal voter's wait is significantly shorter than that of the most negatively affected voters — as seen in the above review, which highlights the large discrepancies between average wait times and extreme long-waits — this treatment is intended to proxy how exposure to long waits impacts on voters' beliefs and support for potential election reform aimed at curbing wait times. The treatment consists of photographs of long lines to vote featuring a cross-section of American voters (examples of which are included in Figures 5.1 and 5.2 below), and a brief summary of the research detailed in 5.2.[35] The primary hypothesis here is that the informational treatment will increase support for election reform. Note that, in order to control for the potentially confounding effects of COVID-19 (e.g., beliefs that long lines are only a product of pandemic-specific constraints), all information treatments include photos that depict voters both wearing and not wearing masks while queuing.

[35] Full text of the information provided in the treatment conditions is available in Appendix A.5.8.

Figure 5.1: Example photograph for T1 (pre-COVID)



Figure 5.2: Example photograph for T1 (COVID)



T2 & T3: Information Treatment + Identity

This portion of the experiment combines the information treatment from T1 with a design that emphasizes the identity of inconvenienced voters. Based on past research that shows how racial imagery affects Americans' support of or opposition to VID laws (Wilson, Brewer & Rosenbluth, 2014), here I seek to determine whether and how race influences views about voting access and reform. T2 uses images that prominently highlight Black voters negatively affected by long waits, while T3 does the same for white voters (examples in Figures 5.3 and 5.4 below).

On the basis of research regarding VID, the hypothesis here is that the identity of inconvenienced voters will affect participants' beliefs about wait times and support for reform. To evaluate this hypothesis, I compare the results of T2 and T3 to the primary information treatment T1, which as mentioned uses photos depicting a cross-section of voter identities. Statistically significant differences between conditions here would suggest that identity does affect the relationship between the information treatment and beliefs, while the lack of such differences would indicate that identity is not a major factor in how people judge the acceptability of long wait times to vote.

Figure 5.3: Example photograph for T2 (pre-COVID)

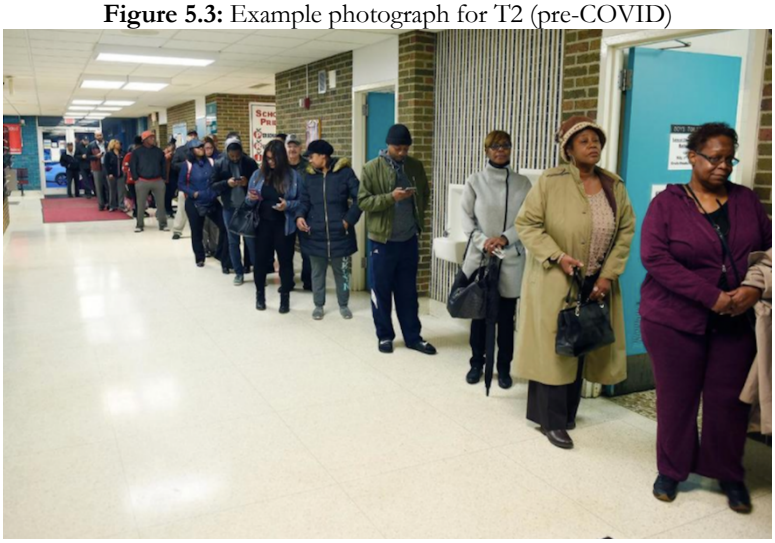


Figure 5.4: Example photograph for T2 (COVID)

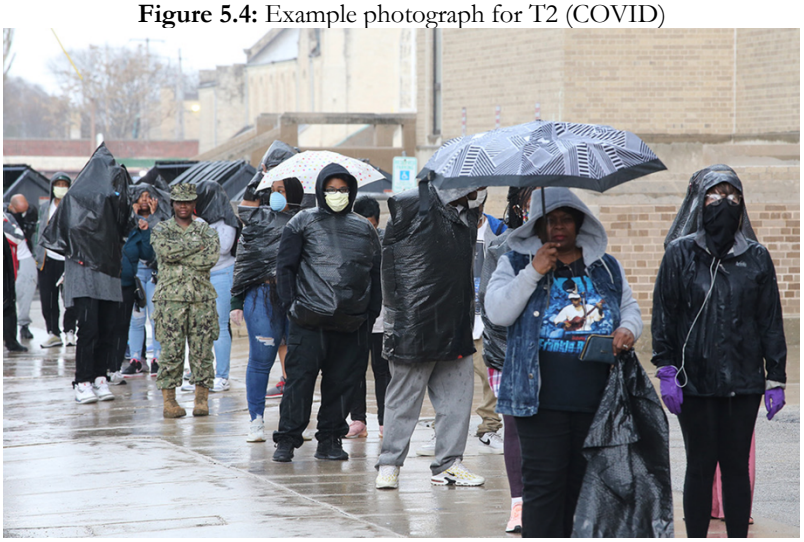


Figure 5.5: Example photograph for T3 (Pre-COVID)



Figure 5.5: Example photograph for T3 (COVID)



T4: Non-US Treatment

Finally, in an attempt to further disentangle the relationship between competitive and democratic preferences, T4 seeks to abstract from the US context to see how personal involvement bears on beliefs about democratic participation. Therein, participants receive an information treatment similar to that in T2, only this time focusing on long wait times in a non-US democracy (Kenya). Participants then answer the same series of outcome questions detailed above, with the main outcome of interest being DV2 (detailed below), about how wait times relate to the health of a democracy. The hypothesis here is that, because of their stake in the outcome of US elections (and lack thereof abroad), respondents may be more likely to use wait times as a proxy of democratic health in non-US contexts.

Figure 5.6: Example photograph for T4 (long waits to vote in Kenya)



5.3.3 Dependent Variables

I employ three outcome questions as dependent variables, gauging the public's view of wait times to vote, which are the means through which I evaluate experimental treatment effects related to the main hypotheses. The primary dependent variable, DV1, is support for reform — of particular interest as it serves as an actionable insight into whether voters might support legislation aimed at getting wait times in line with the Presidential Commission's recommendations. This topic is further probed in the questions described in *3.4d* below, where respondents indicate support of or opposition to a series of policies aimed at curbing wait times. The other two outcome variables of interest gauge how voters believe long waits bear on democratic performance, and the length of wait time that is considered reasonable. The prompts for these DVs are as follows:

> ***DV1***: Generally speaking, would you support or oppose election reform aimed at reducing voting wait times? [0-10 Likert scale from *strongly oppose* to *strongly support*][36]
>
> ***DV2:*** To what extent do you agree or disagree with the statement: *Having to wait a long time to vote is bad for democracy*? [0-10 Likert scale from *strongly disagree* to *strongly agree*]

[36] I opt for the 11-point response scale based on research showing that the validity of Likert response scale monotonically increases in the number of response categories (Lozano, García-Cueto & Muñiz, 2008) and that more scale points enable responses to more closely approach the underlying distribution of opinions (Wu & Leung, 2017).

> ***DV3:*** What do you think is a reasonable amount of time for an eligible, registered, and prepared citizen to have to wait in line to cast their ballot? [numeric-only write-in; time in minutes]

5.3.4 Covariates and Ancillary Questions

Following response on the main outcomes of interest, I include demographic and political questions that are used to evaluate the robustness of experimental treatment effects to factors that might reasonably be assumed to affect responses. The primary covariates are respondents': age, gender, race, political partisanship, education, household income (HHI), and geographical region within the US. These data are used to estimate covariate regressions controlling for important individual characteristics. I further explore potential interaction effects between covariates and treatments, in particular with respect to party affiliation, which has been shown to mediate the relationship between racial imagery and voting beliefs (Wilson et al, 2014).

Beyond these variables that assist in the analysis of the main hypotheses, I further include a number of questions that give context to respondents' beliefs on election administration more broadly. In particular, participants are asked to respond to the following questions:

1) In order to address long wait times to vote, would you support or oppose the following policies? [0-10 Likert from *Strongly oppose* to *Strongly support*]:
    a. Making it easier to vote by mail
    b. Lengthening the period of early voting
    c. Making election day a federal holiday[37]
2) In general, do you support or oppose voter-ID laws? [0-10 Likert from *Strongly oppose* to *Strongly support*]
3) Generally speaking, how much do you trust the American election system? [0-10 Likert from *Not at all* to *A lot*]

The responses to these questions are of particular interest in C1, where they provide insights on the public's views on voting issues without experimental treatment. Moreover, because these questions are, in most cases, well-studied and already the subject of public debate (e.g., voter ID), they further allow us to evaluate the relative fixedness of beliefs on issues for which political

[37] The displayed order of policy proposals is randomised at the individual level.

polarisation already exists. For example, as described previously, in recent years, people's views on mail voting have become crystallised along partisan lines as a result of elite polarisation on the issue (c.f. views on wait times, which may be more amenable to change). If pre-existing partisan divides do indeed strongly affect views on these policy issues, one would expect these questions to be robust to information, and as such to see a null relationship between experimental conditions and responses.

5.3.5 Analysis Plan

The first course of analysis concerns only the control, C1. I consider a number of policy-relevant questions, ranging from how voters feel about long wait times to, more fundamentally, how much trust Americans have in the functioning of their national elections. For these questions, I also consider how subgroups vary in their responses: most notably, splitting participants by partisan allegiance, but also looking for regional, racial, and other differences in beliefs regarding voting and elections. Thus I present full sample means for C1 on the DVs, covariates, and other voting-related questions, as well as — where responses differ — a range of responses broken down along partisan and demographic lines.

The primary analyses then focus on the effect of the experimental treatments on beliefs. Particularly, I am interested in how information affects voting and election-relevant policy beliefs — and if and how partisanship bears on the effect of information. As explained above, T2, T3, and T4 evaluate whether the effect of informing participants about impediments to voting changes depending on the identity of those who are impeded.

To evaluate the hypotheses detailed previously, I compare means for each of the main DVs by experimental condition. I then estimate linear regression models with experimental condition as the primary predictor to check for significant mean differences in results based on treatment.[38] For example, a significant mean difference in support for reform between C1 and T1

[38] With respect to use of a parametric test of independence, I note that there remains methodological debate concerning treating Likert response scales as continuous (Derrick & White, 2017). Recent research suggests that ordinal Likert scores can be treated as continuous, irrespective of the number of categories or the distribution of the

will be used to evaluated H1 — indicating whether or not addressing the information asymmetry between modal voters and the small minority who face long waits affects beliefs about voting wait times and related reforms. Where there are significant differences based on experimental condition, I further estimate linear regression models to include the main demographic and political covariates. Based on previous research, models are also estimated to consider potential interactions between experimental condition and both respondents' race and partisanship.

5.4 Results

5.4.1 Topline Survey Results & Past Voting Experiences

The results of the control yield a number of insights relevant to policymakers and election administrators in the US. Foremost, Americans across the political spectrum largely support reforms aimed at curbing wait times ($\mu = 7.31$) and respondents on the whole also express agreement with the view that long wait times to vote are bad for democracy ($\mu = 6.33$). As for what constitutes a reasonable wait time to vote, the average respondent places this at around 24 minutes, some 20% lower than the 30-minute threshold for wait times recommended by the Presidential Commission. This is an important validation of the Commission's recommendation, showing that public opinion is generally aligned with its view on voting reform. These results are summarised in Table 5.1 below.

Table 5.1: Mean responses on primary outcome variables in C1, full sample

Outcome Variable	μ
Support reform to curb wait times (10 = strongly support)	7.31
Think wait times are bad for democracy (10 = strongly agree)	6.33
Reasonable time for a voter to wait (minutes)	23.58
N	216

variables (Robitzsch, 2020). Reviewing this debate, Harpe (2015) writes that "individual rating items with numerical response formats at least five categories in length may generally be treated as continuous data" (p.842).

As expected, responses to these main outcome variables do vary based on individuals' partisan preferences. In particular, Democrats record the most support for reforms aimed at curbing wait times ($\mu = 8.09$), the greatest belief that long wait times are bad for democracy ($\mu = 7.31$), and the briefest average estimate of what constitutes a reasonable time to wait to vote ($\mu = 20.87$). Yet, though partisanship does significantly affect one's support for policies that would seek to put an end to long wait times to vote in national elections ($p = 0.00$), nonetheless both Democrats ($\mu = 8.09$) and Republicans ($\mu = 6.93$) generally support such reform. Interestingly, Republicans ($\mu = 21.0$) and Democrats ($\mu = 20.87$) agree on what constitutes a reasonable wait time to vote — whereas Independents ($\mu = 28.46$) appear slightly more permissive of longer wait times (though this estimate still falls below the Commission's recommendation). These primary results, split by partisan affiliation, are summarised in Table 5.2 below.

Table 5.2: Mean responses on outcome variables in C1, by partisan preference

	All	**Republicans**	**Democrats**	**Independents**
Support reform to curb wait times (10 = strongly support)	7.31	6.93	8.09	6.83
Think wait times are bad for democracy (10 = strongly agree)	6.33	5.73	7.31	5.83
Reasonable time for a voter to wait (minutes)	23.58	21.00	20.87	28.46
N	216	63	77	76

To investigate whether any of these observed partisan differences in responses to the main outcome variables rise to the level of statistical significance, I estimate a trio of regression models with party ID as the predictor. The results show a small but significant partisan gap in beliefs — with Republicans and Independents clustered more closely together. These are summarised in Table 5.3 below.

Table 5.3: Effect of partisanship on outcome variables in C1

	Support Reform	Reasonable Wait Time	Bad for Democracy
Intercept (= Democratic Party)	8.09 ***	20.87 ***	7.31 ***
	(0.00)	(0.00)	(0.00)
Party ID = Independent	-1.26 **	7.59	-1.48 ***
	(0.00)	(0.31)	(0.00)
Party ID = Republican Party	-1.15 **	0.13	-1.58 ***
	(0.00)	(0.99)	(0.00)
N	216	216	216
R2	0.06	0.01	0.07

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

On the policy variables of interest, I again find a broad range of support for reforms that would make it easier for Americans to vote. Both expanding early voting ($\mu = 7.17$) and making Election Day a federal holiday ($\mu = 7.26$) receive strong and bipartisan support (albeit, with significant differences between parties). Conversely, though expanding access to mail voting is still generally more supported than opposed ($\mu = 6.87$), it is the least popular of the policy proposals and the most polarised by party. Compared to responses on the other two policy proposals, mail voting yields a nearly doubled partisan gap, with Democrats' support moving upward and Republicans' support declining.

This is, perhaps, unsurprising, given previously discussed research on the polarisation of mail voting in the wake 2020 Presidential Election (Lockhart et al, 2020). The observed mean differences on support for mail voting are in line with those observed by Lockhart and co-authors (2020), suggesting that views have begun to crystallise along party lines. This tracks with broader theories on belief polarisation, discussed in the literature review, that show once a policy becomes imbued with partisan affect, partisans tend to take stronger stances in support or

opposition thereof.. Results for each of the policy variables, split by individuals' party preferences, are summarised in Figure 5.8 below.

Figure 5.8: Effect of partisanship on outcome variables in C1



Note: All responses are recorded on a 0-10 scale where 0 = *strongly oppose* and 10 = *strongly support*. The exact wording of the questions is as follows: *In order to address long wait times to vote, would you support or oppose the following policies? 1) Making it easier to vote by mail; 2) Making Election Day a federal holiday; 3) Lengthening the period of early voting.*

The final political and policy questions concern participants' trust in American elections and, relatedly, their priorities for ensuring election integrity. Claims of widespread voter fraud have become ubiquitous amongst certain members of the political class in the US, and there is growing concern that such allegations can undermine trust in elections, and by extension, faith in American democracy (Clayton et al, 2021). Crucially, while asymmetries may exist in the extent to which members of the major political parties promote theories about election manipulation, research has found that both Democrats and Republicans are vulnerable to believing theories about their political rivals engaging in election fraud, and that such beliefs affect individual views of democratic norms (Albertson & Guiler, 2020). This is in line with past research on beliefs

about election integrity in the wake of national elections, which finds that partisans of the losing political party tend to exhibit lower levels of trust in the democratic process in the immediate aftermath of losing an election (Sances & Stewart, 2015). In particular, this relates to a drop in trust that votes have been counted properly — with both Democrats and Republicans more likely to believe in the accuracy of vote tallies when their respective parties have won the election in question (Sances & Stewart, 2015).

Nonetheless, despite some balance in partisan motivated reasoning relating to election integrity depending on whose preferred candidate wins, the last decade has seen the emergence of a significant gap between the major parties regarding general trust in elections and American democracy. Past research has attributed this divide to two main causes: first, following the 2020 Presidential Election, trust in elections increased markedly amongst Democrats, whilst Republicans experienced a slight decrease in confidence; second, even before the 2020 election, this asymmetric trend in partisans' election beliefs had already begun to cement (Clark & Stewart, 2021).

Given these trends, I expect to see a significant divide in participant trust reported based on their party preferences. Table 5.4 below summarises participants' trust in the American election system, with results additionally broken out by political party. In line with previous findings (Clark & Stewart, 2021), the results similar show a partisan gap, with Democrats (μ = 7.69) reporting significantly higher levels of trust than either their Republican (μ = 5.14) or Independent (μ = 5.07) counterparts.

Table 5.4: Mean responses on trust in elections in C1 by partisan preference

	All	**Republicans**	**Democrats**	**Independents**
Generally speaking, how much do you trust the American election system?	6.02	5.14	7.69	5.07
N	216	63	77	76

Note: Responses are recorded on a 0-10 scale where 0 = *not at all* and 10 = *a lot*.

The differences in generalised trust in American elections between parties are highly significant, as can be seen in the regression results shown in Table 5.5 below. A series of additional regression models testing the robustness of this relationship to the inclusion of covariates shows that partisanship remains the greatest and most significant predictor regardless of model specification. Interestingly, age is also a significant predictor of electoral trust, such that older voters (in particular those over 65) report significantly higher confidence in elections than do younger voters, aged 18 to 30 — while neither gender, race, nor region appear to affect trust. The results of these additional covariate analyses are available in Appendix A.5.2.

Table 5.5: Effect of partisanship on electoral trust in C1

	Electoral Trust
Intercept (= Democratic Party)	7.69 ***
	(0.00)
Party ID = Independent	-2.62 ***
	(0.00)
Party ID = Republican Party	-2.55 ***
	(0.00)
N	216
R2	0.18

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

What are the factors that underlie trust in elections? In order to understand voters' priorities when it comes to ensuring election integrity, I pose a pair of questions asking, in essence, what type of error is more acceptable: that is, whether it is more important to prevent ineligible individuals from voting (i.e., prevent illegal voting), or to ensure eligible individuals are able to vote (i.e., ensure legal voting). In general, participants across the political spectrum seem to place

a greater importance on preventing illegal voting than they do ensuring that every eligible voter has the opportunity to cast their vote. Interestingly, there is considerably less variation in these results by party, with no significant differences between Democrats and Republicans in their beliefs about the importance of preventing illegal voting ($p = 0.15$) or ensuring legal voting opportunity ($p = 0.58$).

Table 5.6: Mean responses on legal and illegal voting in C1 by partisan preference

	All	Republicans	Democrats	Independents
People who are NOT legally qualified to vote are prevented from casting a ballot	5.68	5.95	5.05	6.08
People who ARE legally qualified and want to vote are able to cast a ballot	4.83	4.76	4.39	5.34
N	216	63	77	76

Note: Responses are recorded on a 0-10 scale where 0 = *not at all important* and 10 = *very important*. The exact wording of the question is as follows: *Thinking about elections in the US, how important, if at all, is each of the following?*

Finally, I consider how participants' reported past experiences voting correspond to previously discussed research on who faces the longest waits. As expected, there are considerable differences in past exposure to long wait times on the basis of respondents' race, geographic location, and income (though not with respect to partisanship). Note that, because of limited sample size when considering only the control condition, these analyses reflect the experiences of participants across the full sample of respondents. As there are no significant differences in these results by condition (to be expected, given respondents are reporting facts about their past experiences), this allows for a fuller view of people's past experiences with long waits. To focus only on regular voters, these analyses consider only those who voted in the 2020 election ($n = 771$). Tables 5.7 and 5.8 below show the mean longest wait times reported by participants based on their race and region, respectively. As can be seen, Black voters report longer past wait times

than their white counterparts, as do voters in the Midwest and South (both in line with past research).

Table 5.7: Longest previous voting wait time, by race

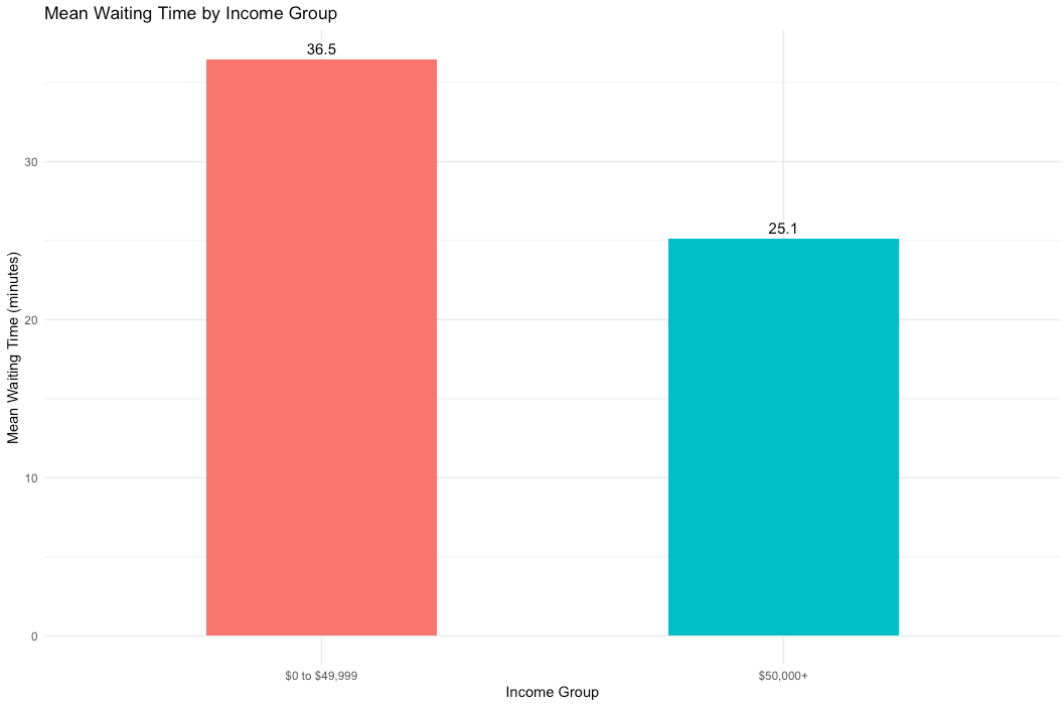
	All voters	White voters	Black voters
Longest past wait time (minutes)	36.1	34.0	45.3
N	771	567	119

Table 5.8: Longest previous voting wait time, by region

	All voters	Northeast	Midwest	South	West
Longest past wait time (minutes)	36.1	28.3	40.6	43.2	28.5
N	771	157	153	274	187

Figure 5.9 below depicts the differences in longest past wait time based on household income, with participants grouped by whether their reported household income is below or above $50,000 per year — the closest approximation of the $48,500 household income threshold used by the Pew Research Center to classify lower income households (Bennett, Fry & Kochhar, 2020). Similar to the trends reported in *3.3b*, respondents from households classified as lower income report longer past wait times than other voters.

Figure 5.9: Longest previous voting wait time, by income (lower income households versus others)



These results provide valuable insight into Americans' beliefs about and experiences with voting and elections. While the analyses do reveal significant differences between partisans' views on a number of crucial issues (foremost, perhaps, election integrity), on the whole responses paint a somewhat optimistic view of Americans' election-related beliefs. Specifically, though the data do confirm other research showing a divide in trust in American elections between the two major parties — with Democrats considerably more trusting than their Republican or Independent co-citizens — nonetheless all three major political groups in the US are on the whole more trusting than untrusting of American elections. Similarly, while I do observe significant partisan differences in support for policies aimed at making it easier to vote in the US, once again all three major political groups are generally supportive of these policies — indicating a unified to desire to make exercising one's right to vote less onerous. Taken together, these are relatively promising results for American democracy.

5.4.2 Information Treatment Versus Control

The first experimental analysis evaluates the effect of the main information treatment — exploring how providing participants with information about long wait times to vote affects their support for making voting easier. In particular, I hypothesised that the information treatment would increase support for reforms aimed at curbing wait times relative to the control. In other words, that individuals' views on long voting wait times and potential policy reforms to address them would be responsive to new information, rather than fixed. I find support for this hypothesis, with a mean increase on support for reform from the control ($\mu = 7.31$) to the treatment ($\mu = 7.96$) of roughly 9% ($p = 0.00$), and a roughly 14% increase in the belief that long waits are bad for democracy ($p = 0.00$). Conversely, views on what constitutes a reasonable wait do not change significantly based on condition, as is generally the case in subsequent treatments. Table 5.9 presents the results of linear regressions for the three main DVs with respondent condition as the sole predictor.

Table 5.9: Linear regression on DV responses, C1 vs T1

	Support reform	Bad for democracy	Reasonable wait
(Intercept = Control)	7.31 ***	6.33 ***	23.58 ***
	(0.00)	(0.00)	(0.00)
Information treatment	0.65 **	0.88 **	-0.47
	(0.00)	(0.00)	(0.90)
N	403	403	403
R2	0.02	0.02	0.00

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

Given previously observed partisan differences, I further consider if and how party preference may mediate the effect of information on beliefs. Table 5.10 compares mean support for reform between the control and information treatment by partisanship. Interestingly, Independents are

the most responsive to the information treatment, with an increase of roughly 13%, while Republicans are the least responsive with an increase of only 2%.

Table 5.10: Mean responses on support for reform, C1 vs T1, by partisan preference

	All	**Republicans**	**Democrats**	**Independents**
Control	7.31	6.94	8.09	6.83
Treatment 1	7.96	7.09	8.49	7.71
N	403	106	170	127

To further interrogate the effects of respondents' characteristics on propensity for policy-related belief change in response to the treatment, I estimate a series of linear regression models (5.11) including key covariates. Model 1 is the most parsimonious of the three, including only experimental condition, gender, age, and geographical region. Model 2 uses the same covariates as Model 1 and further includes political partisanship — while Model 3 adds participant race.

Across all models, condition remains a positive, significant predictor, indicating that the information treatment increases support for reform by between 7 and 10%. As expected, partisanship is also a significant predictor, with Republicans ($p = 0.00$) and Independents ($p = 0.00$) remaining considerably less likely than their Democratic counterparts to support reforms aimed at curbing long waits. Respondents aged 30 to 50 are the most supportive of reform, while neither gender, geographical region within the US, nor race have a significant effect on support for reform.

Table 5.11: Covariate regression models on support for reform, C1 vs T1

	Model 1	Model 2	Model 3
(Intercept = Control)	6.60 ***	7.19 ***	7.24 ***
	(0.00)	(0.00)	(0.00)
Information treatment	0.67 **	0.50 *	0.50 *

	(0.00)	(0.02)	(0.02)
Female	0.10	0.20	0.19
	(0.65)	(0.36)	(0.39)
Age = 30-50	0.75 *	0.63 *	0.65 *
	(0.01)	(0.03)	(0.03)
Age = 50-65	0.32	0.37	0.39
	(0.31)	(0.25)	(0.23)
Age = 65+	0.52	0.56	0.60
	(0.14)	(0.11)	(0.10)
Region = Midwest	0.04	0.24	0.25
	(0.90)	(0.48)	(0.46)
Region = South	0.22	0.35	0.35
	(0.48)	(0.25)	(0.25)
Region = West	0.55	0.63	0.62
	(0.10)	(0.05)	(0.05)
Party ID = Independent		-0.97 ***	-0.97 ***
		(0.00)	(0.00)
Party ID = Republican Party		-1.29 ***	-1.28 ***
		(0.00)	(0.00)
Race = White			-0.09
			(0.73)
N	403	403	403
R2	0.05	0.11	0.11

Note: Gender is coded as: 1 = male, 2 = female. Region is coded as: 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Given the partisan differences in belief change shown above, I also estimate a linear regression model on support for reform to include an interaction between condition and partisanship. Interestingly, though there are partisan gaps in support, there is no significant interaction

between party and condition — indicating that all participants are similarly inclined to adjust their beliefs according to the information received. Because the resulting interaction term is insignificant for respondents of all parties, I have excluded this model from the main text (though it can be found in Appendix A.5.3.)

As for the belief that long wait times are bad for democracy, the effect of information is similarly robust to the inclusion of covariates. Following the same procedure as above, Table 5.12 below summarises covariate regression results for this belief. Depending on model specification, the effect of the treatment is a roughly 12 to 15% increase in the view that wait times bear negatively on democratic performance. Additionally, I again test for an interaction between condition and partisanship, but as the result yields nothing of significance, this model is excluded from the main text and available in A.5.4.

Table 5.12: Covariate regression models on bad for democracy, C1 vs T1

	Model 1	Model 2	Model 3
(Intercept = Control)	5.95 ***	6.52 ***	6.28 ***
	(0.00)	(0.00)	(0.00)
Information treatment	0.92 ***	0.75 **	0.77 **
	(0.00)	(0.01)	(0.00)
Female	0.17	0.28	0.33
	(0.53)	(0.30)	(0.23)
Age = 30-50	1.04 **	0.97 **	0.92 *
	(0.00)	(0.01)	(0.01)
Age = 50-65	0.64	0.76	0.69
	(0.11)	(0.06)	(0.09)
Age = 65+	0.92 *	1.03 *	0.88
	(0.03)	(0.02)	(0.05)
Region = Midwest	-0.77	-0.54	-0.60

	(0.08)	(0.20)	(0.16)
Region = South	-0.67	-0.53	-0.54
	(0.08)	(0.16)	(0.15)
Region = West	-0.05	0.04	0.06
	(0.91)	(0.92)	(0.89)
Party ID = Independent		-0.91 **	-0.90 **
		(0.01)	(0.01)
Party ID = Republican Party		-1.49 ***	-1.54 ***
		(0.00)	(0.00)
Race = White			0.39
			(0.22)
N	403	403	403
R2	0.06	0.11	0.11

Note: Gender is coded as: 1 = male, 2 = female. Region is coded as: 1 = Northeast, 2 = Midwest, 3 = South, 4 = West. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Finally, with respect to what Americans consider a reasonable time to wait, there are no significant differences between the control ($\mu = 23.58$) and treatment ($\mu = 23.11$). As this is true across conditions (with the exception of the non-US treatment, T4), I conclude that respondents have a relatively fixed conception of what constitutes a reasonable wait time, likely based on their own experiences voting.

5.4.3 Identity-based Treatment Effects

We have seen so far that people's views about wait times are dynamic to information, such that bridging the information gap about long waits significantly increases support for policy reform. Does the identity of the affected voters matter? As detailed above, prior research finds that racial imagery, particularly in media portrayals of issues like VID laws, alters the relationship between information and other voting-related beliefs. As such, H2 tests whether the effect of the

information treatment varies when the photographs used prominently highlight Black voters waiting to vote.

As can be seen in Table 5.13 below, compared to the main information treatment T1, participants who are treated with the same information about the prevalence of long waits and their pernicious effects — but who see only photos of Black voters waiting to vote —are significantly less likely to support reform. This drop in support between the treatments is such that it counteracts the previously positive effect of information altogether.

Table 5.13: Linear regression on support for reform by condition (T1 vs. T2)

	Model 1
(Intercept = Information treatment)	7.96 ***
	(0.00)
Information, Black voters	-0.56 *
	(0.02)
N	379
R2	0.02

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

However, further analysis shows this change is sensitive to the inclusion of covariates. Table 5.14 below shows the same comparison of support for reform between the simple information treatment T1 and the information treatment depicting Black voters, T2, with Model 1 adding age, gender, and region, Model 2 further adding race, and Model 3 including each of these covariates plus an interaction of condition and political party. While the sign of the coefficient for Black voters remains negative, Model 2 sees this effect fall below the threshold of significance ($p = 0.09$). Once again, though partisanship remains a significant predictor of participants' views, there is no significant interaction between party preference and condition. Nonetheless, it is noteworthy that, when compared to the control, the information treatment depicting Black voters yields no significant differences in support for reform with or without the inclusion of covariates and a condition-partisan interaction. Thus, it does seem that, when

participants see only Black voters waiting to vote, the result is to counteract the effect of the information treatment on beliefs. These models are available in A.5.5.

Table 5.14: Tiered covariate models on support for reform by condition (T1 vs. T2)

	Model 1	Model 2	Model 3
(Intercept = Information treatment)	7.73 ***	8.06 ***	8.02 ***
	(0.00)	(0.00)	(0.00)
Information, Black voters	-0.56 *	-0.38	-0.27
	(0.02)	(0.09)	(0.44)
Age = 30-50	0.04	0.10	0.11
	(0.90)	(0.72)	(0.71)
Age = 50-65	-0.12	0.01	0.01
	(0.71)	(0.97)	(0.98)
Age = 65+	-0.04	0.20	0.21
	(0.91)	(0.57)	(0.57)
Female	-0.02	-0.09	-0.10
	(0.93)	(0.70)	(0.66)
Region = Midwest	0.58	0.97 **	0.98 **
	(0.11)	(0.01)	(0.01)
Region = South	0.36	0.64 *	0.65 *
	(0.28)	(0.05)	(0.04)
Region = West	0.02	0.16	0.16
	(0.96)	(0.66)	(0.65)
Party ID = Independent		-0.86 **	-0.84 *
		(0.00)	(0.03)
Party ID = Republican Party		-1.65 ***	-1.50 ***
		(0.00)	(0.00)
Information, Black voters x Independent			-0.07
			(0.89)

Information, Black voters x Republican Party			-0.30
			(0.59)
N	379	379	379
R2	0.03	0.11	0.11

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

Similarly, though the average response for what constitutes a reasonable wait time in the second treatment is nearly 4 minutes greater than that in the simple information treatment, this result is below the threshold of significance ($p = 0.058$). No significant relationship is observed with respect to the belief that wait times are bad for democracy.

The drop in support for reform when depicting Black voters is the only significant difference between any of the information treatments and the primary treatment. Table 5.15 provides a full summary of the three identity-based treatments, including covariates for partisanship, race, as well as interactions between these variables and condition.

Table 5.15: Tiered covariate and interaction models on support for reform by condition (all treatments)

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept = Information treatment)	7.96 ***	8.49 ***	8.49 ***	7.97 ***	7.86 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Information, Black voters	-0.56 *	-0.40	-0.36	-0.56 *	-0.38
	(0.02)	(0.08)	(0.30)	(0.02)	(0.36)
Information, White voters	-0.23	-0.11	-0.11	-0.23	-0.10
	(0.30)	(0.61)	(0.73)	(0.30)	(0.81)
Non-US treatment	-0.04	0.04	-0.01	-0.04	0.05
	(0.86)	(0.86)	(0.98)	(0.86)	(0.90)
Party ID = Independent		-0.86 ***	-0.79 *		
		(0.00)	(0.04)		

Party ID = Republican Party		-1.29 ***	-1.40 ***		
		(0.00)	(0.00)		
Black voters x Independent			0.03		
			(0.96)		
White voters x Independent			-0.07		
			(0.89)		
Non-US x Independent			-0.28		
			(0.60)		
Black voters x Republican Party			-0.15		
			(0.79)		
White voters x Republican Party			0.10		
			(0.85)		
Non-US x Republican Party			0.47		
			(0.40)		
Race = White				-0.01	0.14
				(0.94)	(0.70)
Black voters x white					-0.25
					(0.62)
White voters x white					-0.20
					(0.68)
Non-US x white					-0.14
					(0.79)
N	772	772	772	772	772
R2	0.01	0.07	0.07	0.01	0.01

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

As for the comparison of the primary information treatment and the second identity treatment depicting only white voters queuing, though there are small differences there between, I find no

statistically significant differences between these conditions for any of the DVs. This is perhaps unsurprising, as the demographic composition of photos in the primary treatment are such that (as is the case with US demographics) the majority of voters depicted are white, and past research has only indicated a one-way effect of race. Whereas the information treatment featuring Black voters counteracted the effect of the information treatment on support for reform, responses in the second identity condition largely mirror those in the primary treatment for all DVs. Models comparing responses on each of the primary DVs between T1 and T3, with and without covariates, are available in A.5.6 and A.5.7, respectively.

With respect to the non-US treatment, T4, the only statistically significant effect of the treatment is to raise mean perceptions of a reasonable wait time relative to the primary information treatment. As for the hypothesis that changing the competitive context would increase belief that wait times are bad for democracy, while the treatment does yield a small directional change, this change is insignificant regardless of model specification. Thus, I reject the hypothesis that democratic preferences are conditional on competitive interest.

On the whole, then, there is only moderate evidence that contextual factors moderate the relationship between the information treatment and respondents' beliefs. While the comparisons between the identity treatment depicting Black voters and the control do suggest that seeing only Black voters affected by long waits counteracts the effectiveness of the information treatment, the comparisons between treatments show that this change is on the margins of statistical significance, and is sensitive to model specification.

5.5 Discussion

This research is the first of its kind to focus on people's beliefs about impediments to democratic participation at the point of voting. As such, the results of the control alone provide a valuable resource for academics and policymakers alike to understand how people believe long waits bear on the health of American democracy. The results of subsequent experimental conditions make two primary contributions to literatures on beliefs and belief change. First, across political parties, people's policy preferences on issues — at least those that are not previously sapped with the negative effects of partisan polarisation, particularly in the form of elite cues as to what they should believe — are quite responsive to new information. This stands in contrast to related views, such as those on mail voting, for which prior polarisation undercuts the propensity for belief change. The results of the primary information treatment demonstrate that, where information asymmetries may exist, increased information can translate to meaningful and bipartisan increases in support for policy reform. Second, though there are some differences in people's views about the time-tax on voting based on who is depicted as suffering the consequences of these impediments, racial discrimination does not bear as heavily on beliefs and belief change as it has been shown to in other voting-related contexts, such as support for voter-ID restrictions.

There are, nonetheless, some shortcomings in my design and a number of ways on which it could be improved in subsequent research. Most notably, my attempt to evaluate the impact of competitive preferences on perceptions of the relationship between wait times and democratic health in the non-US treatment falls short of providing meaningful insight. While this may be in part due to the lack of more similar cases for comparison — put simply, these waits do not exist in other highly-developed democracies — this underscores the demand for a reconceptualization of this question that may yield more actionable insight. In light of this, a fruitful direction for future research would be to explore whether adding a strong partisan cue attenuates the effect of the information treatment on beliefs. One particularly relevant way to approach this would be to

consider how the inclusion of information about the closeness of the election in which voters suffered long waits bears on support for reform, or even to consider this in the context of a presently-occurring election (for example, long waits for early voting that could conceivably affect the outcome of a current race). Similarly, though the fact that related views on making it easier to vote by mail do not change — despite the obvious reality that so doing would attenuate the presence and impact of long queues — suggests that previous polarisation undermines propensity for belief change, an improved design might explicitly test this hypothesis. One way to do so would be to elucidate in a subsequent information treatment the direct link between mail voting and wait time reduction, and examine whether these views become more dynamic to information.

On the whole, I believe these results reflect rather well on people's willingness to adapt their policy views to new information. In particular, that addressing information asymmetries may be a fruitful way of fostering more agreement on substantive policy reforms. Nonetheless, it is my view that this lesson is unlikely to extend to issues that are already polarised, and for which elite actors and notable politicians have provided unambiguous signals on what partisans should believe. Though the results to this end are only suggestive, and would be bolstered by a design that incorporates the above alterations, the framework of belief responsiveness observed in this experiment is aligned with the notion that previous consideration or polarisation (or the lack thereof) of a political or policy issue powerfully shapes people's propensity for reconsidering or maintaining their beliefs.

Finally, notwithstanding potential threats to American democracy, my results suggest that most people still generally trust elections and election results, and those who do not are but a minority whose noise exaggerates the extent of underlying agreement therewith. In the present context of intense polarisation and the active undermining of democratic principles by some powerful actors, I take this as good news about how likely we are to see elections — the central function of democracy — dramatically destabilised by poisonous and unfounded rhetoric.

5.6 References

Albertson, B., & Guiler, K. (2020). Conspiracy theories, election rigging, and support for democratic norms. *Research & Politics*, *7*(3), 1-9.

Alvarez, R. M., Ansolabehere, S., Berinsky, A., Lenz, G., Stewart III, C., & Hall, T. (2009). 2008 survey of the performance of American elections. Caltech/MIT Voting Technology Project, http://vote.caltech.edu/drupal/node/231

Alvarez, R. M., Hall, T. E., Levin, I., & Stewart III, C. (2011). Voter opinions about election reform: Do they support making voting more convenient?. *Election Law Journal, 10*(2), 73-87.

Alvarez, R. M., Hall, T. E., & Llewellyn, M. H. (2008). Are Americans confident their ballots are counted?. *The Journal of Politics*, *70*(3), 754-766.

Atkeson, L. R., & Saunders, K. L. (2007). The effect of election administration on voter confidence: A local matter?. *PS: Political Science & Politics*, *40*(4), 655-660.

Banks, A. J., & Hicks, H. M. (2016). Fear and implicit racism: Whites' support for voter ID laws. *Political Psychology, 37*(5), 641-658.

Belenky, A. S., & Larson, R. C. (2006). To queue or not to queue. *OR/MS Today*, *33*, 30-34.

Bennett, J., Fry, R., & Kochhar, R. (2020, July 23). *Are you in the American middle class? Find out with our income calculator*. Pew Research Center. https://www.pewresearch.org/fact-tank/2020/07/23/are-you-in-the-american-middle-class/

Bowler, S., Brunell, T., Donovan, T., & Gronke, P. (2015). Election administration and perceptions of fair elections. *Electoral Studies*, *38*, 1-9.

Burdein, I., Lodge, M., & Taber, C. (2006). Experiments on the automaticity of political beliefs and attitudes. *Political Psychology*, *27*(3), 359-371.

Chen, M. K., Haggag, K., Pope, D. G., & Rohla, R. (2022). Racial disparities in voting wait times: evidence from smartphone data. *Review of Economics and Statistics*, *104*(6), 1341-1350.

Ciuk, D. J., & Yost, B. A. (2016). The effects of issue salience, elite influence, and policy content on public opinion. *Political Communication, 33*(2), 328-345.

Clark, J., & Stewart III, C. (2021, April). The Confidence Earthquake: Seismic Shifts in Trust and Reform Sentiments in the 2020 Election. In *annual meeting of the Midwest Political Science Association*.

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.

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, *6*(1), 2053168018822174.

De Kadt, D. (2017). Voting then, voting now: The long-term consequences of participation in South Africa's first democratic election. *The Journal of Politics*, *79*(2), 670-687.

Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics, 18*(3).

Druckman, J. N., & Nelson, K. R. (2003). Framing and deliberation: How citizens' conversations limit elite influence. *American Journal of Political Science, 47*(4), 729-745.

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review, 107*(1), 57-79.

Fortier, J. C., Stewart III, C., Pettigrew, S., Matthew, W., & Harper, T. (2018). Improving the Voting Experience. *Washington, DC: Bipartisan Policy Center*.

Gaines, B. J., Kuklinski, J. H., Quirk, P. J., Peyton, B., & Verkuilen, J. (2007). Same facts, different interpretations: Partisan motivation and opinion on Iraq. *The Journal of Politics, 69*(4), 957-974.

Gilliam Jr, F. D., & Iyengar, S. (2000). Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science*, *44*(3), 560-573.

Gimpel, J. G., Dyck, J. J., & Shaw, D. R. (2006). Location, knowledge and time pressures in the spatial structure of convenience voting. *Electoral Studies*, *25*(1), 35-58.

Gimpel, J. G., & Schuknecht, J. E. (2003). Political participation and the accessibility of the ballot box. *Political Geography*, *22*(5), 471-488.

Gomez, V., & Doherty, C. (2021, July 21). Wide partisan divide on whether voting is a fundamental right or a privilege with responsibilities. Pew Research Center.

Harpe, S. E. (2015). How to analyze Likert and other rating scale data. *Currents in Pharmacy Teaching and Learning, 7*(6), 836-850.

Hicks, W. D., McKee, S. C., Sellers, M. D., & Smith, D. A. (2015). A principle or a strategy? Voter identification laws and partisan competition in the American States. *Political Research Quarterly, 68*(1), 18-33.

Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *The Journal of Politics, 74*(3), 672-684.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.

Kane, J. V. (2017). Why can't we agree on ID? Partisanship, perceptions of fraud, and public support for voter identification laws. *Public Opinion Quarterly*, *81*(4), 943-955.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Levendusky, M. S. (2010). Clearer cues, more consistent voters: A benefit of elite polarization. *Political Behavior*, *32*, 111-131.

Lockhart, M., Hill, S. J., Merolla, J., Romero, M., & Kousser, T. (2020). America's electorate is increasingly polarized along partisan lines about voting by mail during the COVID-19 crisis. *Proceedings of the National Academy of Sciences*, *117*(40), 24640-24642.

Lodge, M., & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology*, *26*(3), 455-482.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73-79.

Mukherjee, E. (2009). Abolishing the time tax on voting. *Notre Dame Law Review*, 85, 177.

Mullinix, K. J. (2016). Partisanship and preference formation: Competing motivations, elite polarization, and issue importance. *Political Behavior, 38*(2), 383-411.

Pettigrew, S. (2017). The racial gap in wait times: why minority precincts are underserved by local election officials. *Political Science Quarterly*, *132*(3), 527-547

Pettigrew, S. (2021). The downstream consequences of long waits: How lines at the precinct depress future turnout. *Electoral Studies, 71*, 102188.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.

Powell, M., & Slevin, P. (2004, December 15). Several Factors Contributed to "Lost" Voters in Ohio. *Washington Post*. https://www.washingtonpost.com/archive/politics/2004/12/15/several-factors-contributed-to-lost-voters-in-ohio/73aefa72-c8e5-4657-9e85-5ec8b2451202/

Presidential Commission on Election Administration. (2014). *The American Voting Experience: Report and Recommendations of the Presidential Commission on Election Administration*. https://www.eac.gov/pcea/pcea

Quealy, K., & Parlapiano, A. (2021, January 4). *Election Day voting in 2020 took longer in America's poorest neighborhoods*. The New York Times. Retrieved from https://www.nytimes.com/interactive/2021/01/04/upshot/voting-wait-times.html

Riker, W. H., & Ordeshook, P. C. (1968). A Theory of the Calculus of Voting. *American Political Science Review*, *62*(1), 25-42.

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. In *Frontiers in Education(5)*, 177-184).

Sances, M. W., & Stewart III, C. (2015). Partisanship and confidence in the vote count: Evidence from US national elections since 2000. *Electoral Studies*, *40*, 176-188.

Scala, D. J., & Johnson, K. M. (2017). Political polarization along the rural-urban continuum? The geography of the presidential vote, 2000–2016. *Annals of the American Academy of Political and Social Science*.

Slothuus, R. (2010). When can political parties lead public opinion? Evidence from a natural experiment. *Political Communication, 27*(2), 158-177.

Spencer, D. M., & Markovits, Z. S. (2010). Long lines at polling stations? Observations from an election day field study. *Election Law Journal*, *9*(1), 3-17.

Stein, R. M., Mann, C., Stewart III, C., Birenbaum, Z., Fung, A., Greenberg, J., ... & Wronski, J. (2020). Waiting to vote in the 2016 presidential election: Evidence from a multi-county study. *Political Research Quarterly*, *73*(2), 439-453.

Stewart, C. (2014). Waiting to Vote in 2012. *Journal of Law and Policy*, 28:439–63.

Stewart, C. (2015). SPAE Report 2015, *2014 Survey of the Performance of American Elections*, Harvard Dataverse V1.

Lodge M., & Taber C. (2000). Three steps toward a theory of motivated political reasoning. In A. Lupia, M. D. McCubbins & S. L. Popkin (Eds.), *Elements of Reason: Cognition, Choice, and the Bounds of Rationality* (pp. 183–213). Cambridge; New York: Cambridge Univ. Press.

Valentino, N. A., & Neuner, F. G. (2017). Why the sky didn't fall: mobilizing anger in reaction to voter ID laws. *Political Psychology, 38*(2), 331-350.

Warren, M. E. (2014). Accountability and democracy. *The Oxford handbook of public accountability*, 39-54.

Westwood, S. J., & Peterson, E. (2020). The inseparability of race and partisanship in the United States. *Political Behavior*, *44*(3), 1-23.

Wilson, D. C., & Brewer, P. R. (2013). The foundations of public opinion on voter ID laws: Political predispositions, racial resentment, and information effects. *Public Opinion Quarterly, 77*(4), 962-984.

Wilson, D. C., Brewer, P. R., & Rosenbluth, P. T. (2014). Racial imagery and support for voter ID laws. *Race and Social Problems*, *6*(4), 365-371.

Wu, H., & Leung, S. O. (2017). Can Likert scales be treated as interval scales?—A Simulation study. *Journal of Social Service Research*, *43*(4), 527-532.

A.5 Appendices

A.5.1 Randomisation checks for participants by condition

Here I test that randomisation has occurred properly by comparing means on demographic variables between experimental conditions. Using age, gender, and respondent region, I observe no significant differences in means between conditions, indicating successful randomisation.

	Age	Gender	Region
Intercept (= Control)	45.40 ***	1.48 ***	2.69 ***
	(0.00)	(0.00)	(0.00)
Information treatment	0.51	0.07	-0.09
	(0.77)	(0.17)	(0.38)
Information, Black voters	-0.63	0.04	-0.07
	(0.72)	(0.37)	(0.50)
Information, White voters	1.64	0.07	-0.01
	(0.34)	(0.14)	(0.95)
Non-US	-0.31	-0.03	-0.10
	(0.86)	(0.60)	(0.34)
N	988	988	988
R2	0.00	0.01	0.00

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

A.5.2 Tiered covariate regressions on electoral trust in C1

	Model 1	Model 2	Model 3	Model 4
(Intercept = Control)	7.69 ***	6.79 ***	7.10 ***	7.33 ***
	(0.00)	(0.00)	(0.00)	(0.00)
Party ID = Independent	-2.62 ***	-2.21 ***	-2.18 ***	-2.18 ***
	(0.00)	(0.00)	(0.00)	(0.00)
Party ID = Republican Party	-2.55 ***	-2.68 ***	-2.64 ***	-2.63 ***
	(0.00)	(0.00)	(0.00)	(0.00)
Female		-0.62	-0.68	-0.67
		(0.08)	(0.06)	(0.06)
Age group = 30-50		1.15 *	1.29 **	1.26 *
		(0.02)	(0.01)	(0.01)
Age group = 50-65		1.23 *	1.41 *	1.40 *
		(0.02)	(0.01)	(0.01)
Age group = 65+		2.38 ***	2.65 ***	2.68 ***
		(0.00)	(0.00)	(0.00)
White			-0.60	-0.61
			(0.17)	(0.17)
Region = Midwest				-0.17
				(0.77)
Region = South				-0.18
				(0.73)
Region = West				-0.44
				(0.40)
N	216	216	216	216
R2	0.18	0.25	0.26	0.26

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

A.5.3 Linear regression on support for reform, C1 vs T1, with partisan x condition interaction

	Model 1
(Intercept = Control)	8.09 ***
	(0.00)
Information treatment	0.40
	(0.22)
Party ID = Independent	-1.26 ***
	(0.00)
Party ID = Republican Party	-1.15 **
	(0.00)
Information treatment x Independent	0.47
	(0.35)
Information treatment x Republican Party	-0.25
	(0.65)
N	403
R2	0.09

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

A.5.4 Linear regression on bad for democracy, C1 vs T1, with partisan x condition interaction

	Model 1
(Intercept = Control)	7.31 ***
	(0.00)
Information treatment	0.40
	(0.34)

Party ID = Independent	-1.48 ***
	(0.00)
Party ID = Republican Party	-1.58 ***
	(0.00)
Information treatment x Independent	0.79
	(0.21)
Information treatment x Republican Party	0.22
	(0.74)
N	403
R2	0.08

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.

A.5.5 Tiered covariate models on support for reform, C1 vs T2, with partisan x condition interaction

	Model 1	Model 2	Model 3
(Intercept = Control)	7.31 ***	7.49 ***	7.50 ***
	(0.00)	(0.00)	(0.00)
Information, Black voters	0.09	0.10	0.14
	(0.71)	(0.67)	(0.73)
Age = 30-50		0.48	0.43
		(0.12)	(0.18)
Age = 50-65		0.36	0.29
		(0.31)	(0.42)
Age = 65+		0.43	0.34
		(0.29)	(0.41)
Female		-0.02	-0.03
		(0.92)	(0.90)
Region = Midwest		0.92 *	0.93 *
		(0.01)	(0.01)
Region = South		0.31	0.32
		(0.36)	(0.34)
Region = West		0.58	0.58
		(0.11)	(0.11)
Party ID = Independent		-1.02 ***	-1.16 **
		(0.00)	(0.00)
Party ID = Republican Party		-1.43 ***	-1.20 **
		(0.00)	(0.00)
Race = White or Caucasian		-0.22	-0.19
		(0.45)	(0.50)
Information, Black voters x Independent			0.30

			(0.61)
Information, Black voters x Republican Party			-0.48
			(0.42)
N	408	408	408
R2	0.00	0.08	0.08

Note: Gender is coded as 1 = male, 2 = female. Region is coded as 1 = Northeast, 2 = Midwest, 3 = South, 4 = West.

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

A.5.6 Linear regression on main DVs, T1 vs T3, no covariates

	Support reform	Bad for democracy	Reasonable wait time
(Intercept = Information treatment)	7.96 ***	7.21 ***	23.11 ***
	(0.00)	(0.00)	(0.00)
Information, White voters	-0.23	-0.04	2.28
	(0.27)	(0.89)	(0.24)
N	389	389	389
R2	0.00	0.00	0.00

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

A.5.7 Covariate regressions on main DVs, T1 vs T3

	Model 1	Model 2	Model 3
(Intercept = Information treatment)	7.60 ***	6.69 ***	20.24 ***
	(0.00)	(0.00)	(0.00)
Information, White voters	-0.12	0.03	2.00
	(0.55)	(0.91)	(0.31)
Age = 30-50	0.38	0.84 *	4.14
	(0.18)	(0.02)	(0.11)

Age = 50-65	0.01	0.26	5.49
	(0.97)	(0.51)	(0.07)
Age = 65+	0.66	0.93 *	2.50
	(0.06)	(0.03)	(0.45)
Female	0.18	0.35	1.90
	(0.38)	(0.18)	(0.33)
Region = Midwest	0.42	-0.24	3.22
	(0.21)	(0.56)	(0.31)
Region = South	0.60 *	0.27	-3.29
	(0.04)	(0.46)	(0.23)
Region = West	0.48	0.42	-0.71
	(0.13)	(0.29)	(0.82)
Party ID = Independent	-0.88 ***	-0.74 *	2.13
	(0.00)	(0.02)	(0.36)
Party ID = Republican Party	-1.41 ***	-1.39 ***	1.96
	(0.00)	(0.00)	(0.42)
Race = White or Caucasian	0.22	0.29	-2.38
	(0.36)	(0.33)	(0.29)
N	389	389	389
R2	0.10	0.08	0.03

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

A.5.8 Full text of information treatment (T1 photographs)

American voters can face long wait times to vote at polling locations across the country.



While the average American waits less than 10 minutes to vote, wait times can be as long as 10 hours, and **roughly 5 million voters waited more than an hour** to cast their ballots in the 2012 election.



Research suggests these wait times affect voter participation, as some voters are unable to spend so much time away from work or home.

For example, MIT's Survey on the Performance of American Elections estimates that **roughly 3 million eligible and registered voters decide not to vote** in an average national election because of prohibitively long lines at polling locations.



These photographs were taken at polling locations across the US in recent national elections (2012, 2016, 2020) and show the length some voters must wait to cast their ballot.



Please review this information and think about your own experiences voting in national elections, about which we'd like to hear on the next page.

6

6 Conclusion

This dissertation began with a question: how do preconceptions affect the evaluation of political information and actors, and how do these evaluations filter through to behaviour? Weaving together evidence on both the social and cognitive nature of prejudicial thinking, I have questioned how we ought to think about the relationship between information, beliefs, and attitudes. That is, does information drive beliefs, bending our views in line with the best available evidence on a given subject, and in turn shaping our attitudes — or, conversely, do our existing beliefs and attitudes colour the way we take in new information, leading us to accept or reject it on the basis of its accordance thereto? If the answer lies somewhere in the middle, such that beliefs and attitudes do bias our consideration of information but not to the extent that evaluations are impervious to change, what factors bear on propensity for belief adjustment?

In the first two chapters, I covered an extensive body of research in political and social psychology on *motivated* thought, the sum of which suggests that — rather than prior beliefs being held outside of the review of new information — existing beliefs and attitudes lead us to selectively seek, injudiciously process, and insufficiently incorporate information into our evaluations. In this framework, individuals are inclined, consciously or unconsciously, to maintain beliefs that align with their broader worldviews. Or, in other words, prior beliefs are the filter through which new information must pass. In this dynamic, information that corresponds to stronger or more salient attitudes is most susceptible to attitude-congruent evaluation — which, according to some renderings, is because issues on which we hold stronger or more salient views, when broached, are more *charged* with previous affective cues that colour the way we perceive new information pertaining thereto. This bias toward accepting attitude-congruent

information effectively amounts to a tax on discordant information: put simply, we need less and less persuasive evidence to arrive at an evaluation that aligns with our beliefs and attitudes, and more and more persuasive evidence to accept an evaluation that contradicts them.

This point was demonstrated, in somewhat distinct manners, in each of the three articles presented herein. Perhaps most directly, the incongruent punishment of political candidates in Article 2 for an ethical transgression can be interpreted, in part, as a testament to the dynamic described above. That is, though people do attempt to suppress or override existing prejudices, information that conforms to a view in line with those prejudices (e.g., stereotypes about Black males and criminality paired with the revelation that a Black male candidate has engaged in unethical behaviour) is reflected more heavily in subsequent evaluations. Conversely, as in the cases of Democrats less severely punishing the Black female candidate and Republicans less severely punishing the white male candidate, the penalty dispensed for the same unethical behaviour is less severe when we are more inclined to hold positive views of that individual because of a perception that they are likely part of our in-group. This understanding of how the strength of past beliefs and attitudes alter the efficacy of information can also be applied to Article 3, wherein the information treatment caused meaningful changes in participants' support for reform aimed at curbing long wait times to vote — a topic, as elaborated, charged with little affect or prior polarisation. Conversely, views about related reforms such as expanding access to voting by mail or making election day a federal holiday (which, though distinct from addressing wait times directly, would nevertheless also make it less costly for people to vote) remained entirely fixed, perhaps in part because of their previous politicisation. Finally, the results in Article 1 present a special case that was particularly likely to engender belief updating, simply because participants could (and for the most part did) update their underlying beliefs about natural and methane gases in line with their broader partisan and issue attitudes. This may help explain why the effect of information on belief convergence observed was so strong: beliefs could update based on new evidence without forcing participants to arrive at a counter-

attitudinal position. Conversely, those who began the experiment with strongest existing beliefs in favour of methane gases were also the least likely to update their misperception on the basis of new information, or in common parlance, to exhibit a backfire effect.

Thus, in each of the three articles that comprise this dissertation, I have touched on only some of the ways in which previously-held views and heuristics bear on important issues in contemporary American politics. Read together, they demonstrate the pervasive effects of motivated and prejudicial reasoning on political judgments, but also some reasons to think these inclinations can be overcome depending on the decision context. A brief review of the main results and lessons from these articles follows.

One Fact Two Beliefs

In Article 1, I tackle head on the question of how responsive people's beliefs are to new information, presenting a novel context in which to evaluate not only whether people update in response to new evidence, but also how these beliefs are likely to change. Reviewing a range of influential papers on factual corrections and belief change, I argue that much previous research fails to effectively delineate between fact and opinion, and that this conflation of the two affects the extent to which people believe there can be an objective answer to the question at hand. As a result, I contend, it seems likely that much past research on corrected misperceptions systematically underestimates people's propensity to change factually-incorrect beliefs in line with new information. To test this, I focus on a widespread misperception that is demonstrably (indeed, as I put it in the article, physically) false, and measure the prevalence, extent, and direction of belief updating in response to the evidence presented.

Therein, I find that, contrary to some depictions of contemporary political actors and debates as fact-resistant or *post-truth*, people across the political spectrum are generally responsive to a factual correction of a previously-held misperception on fossil fuel energy sources. This finding pushes back against the view that factual corrections are ineffective in political contexts,

or — as put forth by the literature on backfire effects — may even engender a retrenchment of inaccurate beliefs. Yet, some of the mechanisms that may underlie such an expectation are present in my results. While people demonstrate no issue with incorporating a belief-disconfirming new fact into their evaluations, it is the way in which this assimilation occurs that speaks more to concerns about polarisation and its resulting impact on the likelihood partisans can reach a point of agreement.

In line with expectations drawn from a rich literature on asymmetric belief change and biased interpretation, detailed both in the article and in Chapter 2, I show that the lessons people take from new information are systematically inclined to favour a view that aligns with their broader partisan and issue attitudes. That is, while the specific misinformed belief is open to changing, the direction of that change is toward an attitude-congruent evaluation. I thus argue that these differences are predictable to the extent that researchers view partisanship not as an active calculation of how well a party platform serves their personal interests, but a facet of individuals' identities and shaper of their underlying evaluations and beliefs. With this novel decision context, I effectively force people — or at least, those who want to align their beliefs with the fact presented — to choose between the options: *this thing I thought was bad is actually good,* or *this thing I thought was good is actually bad*. Much in line with the models of attitude-affirming cognition described in the literature review, when presented with these two ways of changing their beliefs, people are significantly more likely to do so in the direction of their broader partisan and climate attitudes. Moreover, participants who had originally expressed the strongest opinions in favour of methane-based gases were significantly more likely to exhibit a backfire effect, further retrenching their misperception. This, too, follows from the models outlined in the literature review, which argue that the strength of past a belief predict its impact on a subsequent judgment. On this basis, I argue that — to the extent that issue attitudes are previously polarised, as in the case of climate attitudes in the US — belief adjustment is likely to engender more, not less, disagreement.

Does Candidate Identity Affect Voter Response to a Scandal?

The lessons drawn from *One Fact, Two Beliefs* are also one interpretation for the results of *Candidate Identity*. Whereas people who hold prejudice are reluctant to express it for fear of social sanction, these prejudices reveal themselves given an opportunity to ascribe them to other causes. Grounded in past research showing differences between the ways in which prejudice is expressed versus held (which I conceptualise as latent prejudice), I adapt and add to an experimental toolkit for understanding the powerful effect of anti-prejudice norms (or lack thereof) on the suppression (or revelation) of prejudice. To do this, I detail a supposed ethical transgression on the part of potential candidates for state office (intended to proxy the effects of a political scandal), which affords respondents a cause to which to attribute any subsequent negative evaluations. With the changing of the norm, I observe a revelation of prejudice against Black males from respondents across the political spectrum, such that the penalty incurred (calculated, between subjects, as differences in how candidates are rated with and without the scandal) for the transgression significantly exceeds the same measure for peer candidates.

As elaborated above, this, too, is a motivated conclusion: in the absence of evidence to affirm a negative view of a candidate, people may extend to political actors similar levels of openness and receptivity. Yet, the playing field is far from level. Black male political actors are afforded a significantly lower threshold for even small violations of ethical conduct, and punished more severely when their actions stray from an ideal. One way to interpret these findings is that, because of the legacy of learned negative stereotypes against Black males in American culture, people weight individual instances of unsavoury behaviour more heavily in their evaluations — yet, would require more evidence of such ill-conduct to similarly adjust their views toward a white candidate. As detailed in the article, this same pattern seems to hold based on partisan allegiances, as clearer signals that a candidate may be an in-partisan (as in the case of the Black female for Democrats and the white male for Republicans) lead to significantly less

severe penalties for the transgression. Here is the injudicious processing previously alluded to: the same action is more damning when we are already inclined toward a negative view. Again this is in line with the models of information-processing and judgment laid out in Chapter 2 and detailed above: we need less information to arrive at conclusions that are easier for us to digest, and more for those we would rather not arrive at. Or, in other words, information and actions that conform to our prejudicial expectations are more significantly reflected in subsequent evaluations than those that contradict them. In either light, the results presented in this paper powerfully demonstrate the effect of prejudiced thought on our evaluation of political information and actors, and that the legacy of racism in America still bears on our politics.

Give US the Ballot?

If there is a silver-lining in these pages, it is to be found in Article 3, which shows that — on issues that are not already coloured by partisan rhetoric and elite cues — people seem to arrive at similar conclusions about a simple policy issue that, if acted on, would sum to profound changes to the administration of American democracy. Therein, I seek to address an information asymmetry resulting from people's vastly different voting experiences: the very uneven distribution of long wait times to vote, the burden of which fall on only a small segment of the electorate. On the basis of this experiential gap, I posit that Americans' support for policy reform to address long wait times may be receptive to new information. In testing this hypothesis, I find people across the political spectrum to be receptive, such that the information treatment highlighting long waits engenders meaningful changes in their beliefs. This translates to significantly increased support for policies that would make it easier for all Americans to participate in democracy, with none of the partisan differences that characterise other, already polarised issues around voting access and elections. Moreover, while some of the trends surrounding racial prejudice observed in Article 2 are also seen here — with depictions of only Black voters waiting to vote effectively eroding the impact of the treatment — these results are

nowhere near the magnitude observed in past research on voter-ID laws, and suggest that at least in principle people believe all voters, regardless of identity, should have the chance to exercise their right to vote. Relating back to the framework outlined above, these results indicate that, lacking strong motivation to arrive at a particular conclusion, people's political and policy beliefs and attitudes are open to adapting on the basis of new information they deem persuasive.

Concluding Remarks

In light of these findings, what can be learned with respect to the questions postulated at the start of this section? For starters, prejudicial and motivated thinking — whether based on partisan, identity, or issue-based preferences, and whether conscious or unconscious — profoundly impact the ways in which we evaluate and interpret new political information. In the absence of strong existing priors or externally imposed cues for what we should think, our beliefs are open to forming or changing on the basis of new, high-quality information. Thus we are willing to update a misperception or adapt our policy beliefs if so doing does not seem to contradict a view we are inclined to hold. Unfortunately, this narrows the scope considerably for how effective information can be in changing political attitudes and beliefs. Taking the inverse, for salient or strongly held beliefs and attitudes, it seems unlikely that new information alone can effectuate much change.

The sum of this dissertation, then, is much aligned with the understanding of motivated cognition and political judgment elaborated in Chapter 2. Briefly, we are all biased actors. Trying to make sense of and order the vast amounts of information around us, we filter new evidence through our dual lenses of previous belief and unconscious association. The upshot is that that we will draw from new information the conclusions that best suit our preferred view of the world, "to the extent that reason permits" (Kunda, 1990, p.483). In a polarised political context, this means that giving two out-partisans the same information may only serve to exacerbate, rather than alleviate, their disagreement — or, as quoted previously, "the fact that people

interpret information in a self-serving manner means that ... giving two parties more information may cause their expectations to diverge" (Babcock et al, 1995, p.1337). Information, therefore, is no panacea to political polarisation and extremism. Rather, the more polarised or extreme our beliefs are, the less likely information is to effectuate changes to our beliefs and attitudes. If we are to successfully address these issues, we must first acknowledge and seek to understand these realities of human judgment and decision-making.