Defining and Measuring Democratic Norms*

Gretchen Helmke and Josiah Rath

July 29, 2024

Abstract

If scholars and pundits are right, the erosion of norms in the United States and abroad poses a significant danger to democracy. Understanding what exactly norms are, what makes them "democratic," and how best to measure them are thus essential building blocks for generating and evaluating explanations of how such norms weaken and collapse. Our essay addresses each of these key elements. On the conceptual front, we argue for more precision in defining norms and more consideration in labeling them as democratic. On the measurement front, we develop a general utility function and use it evaluate the various methodological strategies that researchers have deployed to causally identify democratic norms. In-between, we synthesize the fast-growing literature on norms and democratic backsliding using a four-fold typology, with transgressors and enforcers on one dimension, and political elites and citizens on the other. We conclude by pinpointing several new areas for future research.

Introduction

Norms have long been a central focus for scholars of democracy. Often used interchangeably with terms like culture, values, and informal institutions, core cooperative norms —think trust and reciprocity— have historically been viewed as part and parcel of self-enforcing democracies. Tocqueville (1840) famously argued that democracy sharpened the

*Currently under review

social mores and sentiments of Americans. Foundational work on comparative politics a century later by scholars such as Lipset (1959) and Almond & Verba (1963) stressed that the values and culture of a country's citizens were crucial for stabilizing democracy. Meanwhile, following the third wave of democratization, the persistence of anti-democratic norms related to clientelism, patrimonialism, and corruption were blamed for many of the failures of new democracies to consolidate (O'Donnell 1996), as well as the inability of certain more advanced democracies to work well (Putnam et al. 1994).

Now, with much of the discipline riveted by the global challenges of democratic backsliding, norms have once again come to the fore. As Figure 1 shows, the number of political science articles including the term "democratic norms" has been steadily increasing among scholars since the 1990s, with a sharp uptick post-2016. In stark contrast to earlier waves of research, however, much of the recent scholarship portrays norms as surprisingly fragile. Indeed, such frailty is perhaps *the* key common feature across the many countries that have experienced backsliding over the last decade. Even as scholars continue to debate the extent to which global democratic backsliding has affected formal institutions (see Little & Meng 2024, and responses in the Special Issue on Democratic Backsliding in PS: Political Science and Politics), few doubt that many basic democratic norms lie in shreds.

Consider the United States, where formal institutional safeguards largely held under the Trump administration, but norms did not.[1] Although Congress, the courts, and bureaucracies mostly managed to constrain the administration's worst impulses, Trump serially broke countless long-standing presidential norms, ranging from failing to release his tax records and lying about the size of his inaugural crowd, to attempting to politicize the military and intelligence agencies, to undermining scientific experts and spreading disinformation during the global pandemic, to repeatedly sowing doubt about the integrity of the vote count and refusing to concede the 2020 election after his definitive loss. Similar stories can be told about Brazil under Jair Bolsonaro, Mexico under López Obrador, and India under Narendra

[1] For a list of the norms violated under Trump, see The Washington Post, "The Abnormal Presidency" Nov 10, 2020.

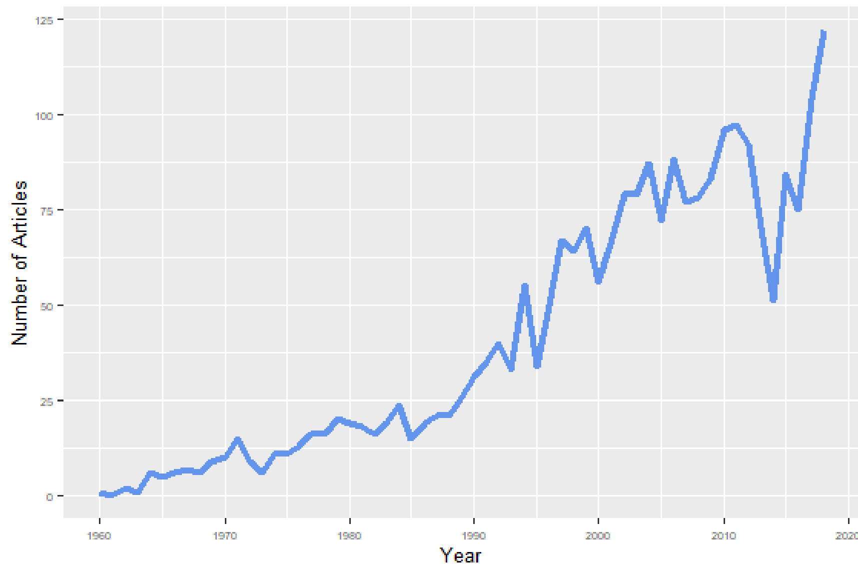

Figure 1: Graph constructed by searching JSTOR for articles in Political Science Journals containing the term "democratic norms" in the text (1960-2018).

Modi.

If scholars and pundits are right, the erosion of norms both in the United States and abroad poses a significant danger to democracy and the rule of law. As Levitsky & Ziblatt (2018) argue in their seminal book *How Democracies Die*, because constitutions are necessarily incomplete contracts, filled with "countless gaps and ambiguities" and "subject to competing interpretations" norms are essential for enabling governments—as well as constraining them. As any student of the American presidency knows, both the power that the executive office enjoys, and the independence that its grants to agencies like the DOJ are not specified in Article II; they are largely based on norms (see Azari & Smith 2012). In the absence of the right norms, the formal rules can easily be subverted; a liberal democracy can unravel by having its own institutions weaponized (Varol 2014; Ginsburg & Huq 2018). Thus, if we are to move forward as a discipline devoted to explaining how democracies thrive or decay, it is crucial to understand what norms are, and what it is that makes a particular norm "democratic."

This is where we begin our essay. Drawing on a broad transdisciplinary literature, we flesh out a definition of norms as informal behavioral rules based on shared expectations

about what others will do, and what others believe we *ought* to do. Norms, in a nutshell, involve preferences that are fundamentally conditional on mutual expectations and, thus, fundamentally social (Bicchieri 2006; 2017). What makes a given norm democratic, we argue, is often less clear. In our view, few norms are inherently pro- or anti-democratic; rather the task of the scholar is to make the case context-by-context using a coherent theory of democracy rooted in a disciplinary consensus (see Przeworski 2024).

The second part of our essay synthesizes the sprawling recent literature on democratic backsliding and norms using a basic four-fold typology. The first dimension reflects the idea that all norms involve an on-going relationship between potential transgressors and potential enforcers. The second dimension captures the democratic context and distinguishes between citizens and political elites. For each permutation of this dyad, we identify the core mechanisms purportedly sustaining norms —accountability, custodialism, self-interest, and conformity—and consider a wide range of explanations for how each might begin to fail.

The third part of the essay shifts to measurement and identification. We introduce a general utility function to capture the idea that the decision to abide by a norm rests on some latent combination of unconditional and conditional preferences. The task of the scholar seeking to discern whether norms are driving behavior, then, is to identify the extent to which decisions hinge on higher-order beliefs (i.e., either first-order "empirical expectations" about what others will do, or second-order "normative expectations" about what others believe one ought to do) versus personal preferences and values. We discuss the advantages of (and limits to) the various research designs that scholars have developed to begin to causally disentangle these different motivations.

We conclude with a call for future researchers to focus on three main areas of inquiry. The first involves exploring how altering whose expectations matter, matters. In much of the applied literature, scholars have tended to focus entirely on gauging the effects of shifting expectations about out-partisans. Manipulating the identity of the reference group, say, by correcting respondent's views of co-partisans, is an obvious next step in the broader

research agenda on democracy and misperceptions. The second area highlights the need for more research on the process of "normalization"—that is, how norm-breaking becomes internalized and legitimated over time. The third and final frontier invites scholars to expand their focus to include not just the democratic norms that have been destroyed, but to identify and understand the new "post-democratic" norms that are being created in their stead.

Conceptualizing Democratic Norms

On the definitional front, political scientists have yet to fully converge on a common understanding of democratic norms. Voluminous bodies of scholarship on the concepts of democracy and norms exist.[2] Yet, many of the careful distinctions made within each body of literature are often strangely absent whenever norms and democracy are considered jointly. In both contemporary popular and academic discourse, democratic norms are largely treated as metaphors—think "rules of the game" and "soft guardrails"—or simply left undefined altogether.

The first goal of this section is to make explicit the criteria by which we distinguish norms from other frequently associated terms, such as attitudes, values, or principles, as well as from the closely related concept of informal institutions. The second goal is to encourage greater reflection on the standards by which a given norm is considered democratic. Our view is that because most norms are not, in fact, inherently democratic or undemocratic, how one characterizes the "democratic-ness" of a norm depends largely on how one characterizes the status quo of the current political regime, which itself depends on how narrowly or broadly one defines democracy.

We begin with norms. A first cut might simply refer to norms as enforced patterns of be-

[2] Beyond the field of political science, norms are foundational in sociology (Parsons 1951; Bourdieu 1984; Coleman 1990), psychology (Cialdini et al. 1991; Sripada & Stich 2007), anthropology (Geertz 1973; Ensminger & Knight 1997; Boyd & Richerson 2005), economics (Schelling 1960; Arrow 1969; Sugden 1989; North 1991; Gintis 2010; Young 2015), philosophy (Hume 1740; Lewis 1969; Ullmann-Margalitt 1977; Bicchieri 2006; Brennan et al. 2013), and law (Sunstein 1996; Ellickson 1998; Posner 2000). For recent reviews of social norms, see Young (2015), Horne & Mollborn (2020), and Gelfand et al. (2024).

havior (Axelrod 1986; Young 2015). Deviations are punished physically (violence, expulsion) or psychologically (guilt, scorn, ostracism, changes in social status). Evolutionary arguments (Axelrod 1986; Skyrms 2004; Bowles & Gintis 2011) contend that behaviors or parameters influencing them (e.g., altruism) are transmitted across generations; successful individuals are more likely to bequeath their traits and behaviors to the next cohort. A second view, pioneered by Schelling's work on focal points, defines a norm as a particular equilibrium to a game with multiple equilibria (Schelling 1960; Ullmann-Margalit 1977; Kandori 1992; Weingast 1997; Mackie 1996). Yet a third approach, developed most fully in economics and philosophy treats norms as rules of behavior embedded with mutually shared expectations (i.e., beliefs) that guide individuals towards an equilibrium (Lewis 1969, Sugden 1989, Elster 1989, and Bicchieri 2006). The latter approach is the one we adopt here.

Specifically, we define political norms as informal behavioral rules that an individual follows in a given political circumstance only if they expect (1) others to follow the rule, and (2) that others believe they ought to follow the rule. The first condition commonly refers to "empirical expectations" and suffices to establish purely descriptive norms. The second condition refers to "normative expectations" and is necessary for creating injunctive or social norms (Bicchieri 2017; also see Goldstein 2024; Dinas et al. 2024).

Distinguishing norm-driven behavior from other motivations thus depends on whether adherence to a given rule is conditional or unconditional. Per Bicchieri, a norm can be said to drive behavior only if the counterfactual is true: Individuals would not continue to follow a given behavioral rule if they did not believe that others also adhere to the rule, or if they ceased to believe that others expected them to follow the rule (Bicchieri 2006:16; 2017:88-93). Conversely, we might say that individuals who follow some behavioral regularity on the basis of principle or habit have a strictly dominant strategy: they abide by a given rule regardless of their expectations of others or others' expectations of them.

To give an example from the political realm, our definition implies that theories of voter turn-out based purely on the so-called "duty" term, are not necessarily norm-based. As

Gerber et al. (2008) reflect, voting only becomes a norm to the extent that is motivated by the desire to avoid disapproval by others. Or, to further illustrate the distinction between norms as expectations and personal values, consider the so-called "master" democratic norm of forbearance described in Levitsky & Ziblatt (2018). As we discuss below, a party's willingness to refrain from fully maximizing formal legal institutional prerogatives against their opponents depends fundamentally on its expectations about whether the opposition also has incentives to pull its punches (cf. Helmke et al. 2021). If a politician nevertheless still puts "country above party," despite believing that her opponent will not do the same, she is not following an existing norm; she is acting on the basis of principle.

For any norm to exist, a sufficient number of people must also recognize that a given circumstance calls for a given behavioral response, or what Bicchieri refers to as the "contingency condition," (Bicchieri 2006:11-13). Just as crucially, we argue, is whether a sufficient number of people also believe that the circumstance itself exists. Within the political realm, norms become especially vulnerable whenever this latter condition is undermined. Divergent perceptions of the 2020 U.S elections offer a particularly vivid example of this problem. Survey evidence shows Republicans overwhelmingly support the norm of concession in the abstract, but because large majorities on the Right do not believe that Biden was the legitimate winner, the norm of concession is inoperative, or at least "trumped" by the norm of supporting free and fair elections (Bright Line Watch 2019; 2020). In other words, even in situations where there remains general agreement over a norm in principle (i.e., Bicchieri's contingency condition), if there is no common ground truth about events, the norm cannot be enforced.

By emphasizing the role of shared beliefs and mutual expectations in shaping informal rules of behavior, our definition is largely consistent, though not necessarily synonymous, with previous scholarship on informal institutions. Helmke & Levitsky (2006:1-88), for example, define informal institutions as "rules and procedures that are created, communicated, and enforced outside of officially sanctioned channels." Our definition of norms encompasses

this understanding, but is broader in two important respects. First, norms need not be sustained purely by an enforcement mechanism that relies on negative sanctions. To be sure, individuals may often follow norms to avoid being punished, but may also do so out of a basic desire for conformity with a particular group, or through internalizing the norm. Second, our definition does not necessarily require that enforcement occurs outside of officially sanctioned channels created by the state. For example, if voters punish politicians who lie or refuse to promise to concede if they lose, then the norm itself is still informal (i.e., there is no law or constitutional provision of which we are aware that requires politicians to tell the truth or gracefully accept a loss), but it is being enforced by the formal institution of voting (see Stokes 2006).

Having roughly pinned down what we mean by norms, what makes a given norm "democratic"? Clearly, not all political norms that exist in a democratic regime qualify. Authoritarian politicians, elected by "people voting against democracy," are given a mandate to undermine it (Przeworski 2024). Likewise, not all norms that are considered democratic in one context are so in another. As Little & Meng (2024) explicate, the norm of conceding electoral defeat should hardly be considered democratic in a context marked by fraudulent elections.

Indeed, to the extent that no democracy is itself ever fully democratic (Dahl 1971), political norms that exist within it necessarily play a double-edged role. That is, norms operate in ways that simultaneously preserve what is democratic in a regime, and what is not. Throughout much of U.S. history, the norm of elite forbearance may well have served to reinforce inter-party contestation and the peaceful alternation of power, but it did so at the heavy price of racial exclusion (Mickey 2015; Levitsky & Ziblatt 2018; Mettler & Lieberman 2020). In other contexts, norms of political forbearance have entailed other important tradeoffs, such as the unequal application of the rule of law (Brinks 2006; Holland 2017; Helmke et al. 2021). Recognizing these sorts of ambiguities inherent in most democratic norms provides an important counterpoint to the sort of nostalgia that often tinges accounts

of democratic erosion.

Ultimately, our point is that classifying norms as democratic should not be done reflexively. It may well be the case that any number of political behaviors loosely considered "normal" under, say, a previous government not engaged in democratic backsliding, are crucial for maintaining democratic stability. The rub, though, lies in specifying exactly how and why a particular norm has functioned as a net benefit for democracy. Just as the noun "democracy' should not be used to refer to all normatively attractive elements of political life (cf. Przeworski et al. 2000), nor should the adjective suffer the same fate.

Transgression and Enforcement: A Typology of Democratic Norms

To organize the fast-growing literature on the role of norms in democratic backsliding, we develop a typology based on two dimensions. The first dimension builds on the basic understanding of norms as inherently revolving around two types of actors: a potential transgressor and a potential enforcer. Within democracies, we can then further distinguish between whether transgressors are theorized as political elites or as citizens, and whether enforcers are envisioned as citizens or as political elites.

This generates four permutations of the transgressor-enforcer dyad. The off-diagonal corners (bottom-left and top-right) of Table 1 correspond to the two possible types of vertical relationships involving citizens monitoring politicians and vice versa, whereas the on-diagonal cells refer to distinct types of horizontal relationships among political elites and across citizens, respectively.

The contemporary literature on norms and democratic backsliding, we hasten to add, is by no means balanced across the cells. Most of the scholarship has either been centered on understanding citizens' failure to hold norm-breaking politicians accountable (the top-right), or on the collapse of forbearance among political elites (the top-left). More recently, scholars

		Enforcer	
		Politicians	Citizens
Transgressor	Politicians	Self-Interest	Accountability
	Citizens	Custodialism	Conformity

Table 1: A 2x2 typology of Democratic Norms between Transgressors and Enforcers

have also begun to examine how democratic norms degrade among citizens; only a handful of studies systematically address how political elites might encourage the preservation of democratic norms among citizens.

Within each dyad we identify the central mechanism by which the enforcement of norms are presumed to operate (or, once backsliding begins, fail to operate). The citizen-to-politician relationship is characterized by *Accountability*, whereby politicians may be tempted to violate democratic norms, but obey them if citizens can credibly threaten to sanction them. Conversely, the politician-to-citizen cell is labeled as *Custodialism* and refers to a scenario more commonly described in earlier waves of literature on democracy, whereby politicians keep citizens in line by modeling appropriate virtues. Variations of *Self-Interest*, in turn, characterizes most explanations of why political elites either opt to follow cooperative norms, or violate them. Finally, *Conformity* serves as the central mechanism purportedly sustaining democratic norms among citizens.

Citizens to Politicians

Much of the research on democratic backsliding begins with the central puzzle of why citizens who routinely claim to value democracy, often fail to punish politicians who violate basic democratic norms. Explanations tend to fall roughly into four groups: trade-offs, cognitive biases, lack of information (by citizens about politicians), and misperceptions (by citizens of other citizens).

The idea that voters care about democracy, but not enough to protect it lies at the heart of Svolik & Graham's (2020) landmark study. The more polarized voters and candidates become, so their argument goes, the more voters will tolerate co-partisan candidates who violate democratic norms (Bartels 2020; Ahmed 2023; Carey et al. 2022). Other scholars have identified a kind of "democratic hypocrisy" whereby citizens are only willing to protect norms when their own party is effectively shut out of power (Simonovitz et al. 2022). Kingzette et al (2021) further establish that affective polarization (i.e., how warmly one feels towards one's own party and how cold one feels towards the out-party) systematically reduces support for democratic norms, particularly those related to forbearance among in-party voters.

Alternatively, citizens may be engaged in a process of "democratic rationalization," in which they interpret similar actions taken by politicians differently depending on whether they share the same ideology as the politician. Krishnarajan (2022), for example, develops the novel argument that whenever norm violations clash with voters' policy preferences, voters will seek to reduce cognitive dissonance either by transmitting policy agreement into support for norm breaking behavior and/or by justifying any violations of norms by politicians with similar ideological views as necessary measures for the broader good of the country. Precisely why the desire to reduce dissonance necessarily leads to individuals' altering their views on norms versus policy, however, is not entirely clear. Notice as well that across these explanations, politicians' willingness to follow norms is presumably conditional, but voters' decisions remain predicated on their sincere preferences or values —however rational or irrational those preferences may be.

Citizens' expectations begin to enter the picture in theories that highlight voter uncertainty. Building on the observation that backsliding is an inherently gradual — and often quite opaque — process (Bermeo 2016; Varol 2014; Levitsky & Ziblatt 2018), for example, Chiopris et al. (n.d.) argue that the motivations of politicians' attacks on democracy may be unclear to voters at first, particularly if politicians are able to effectively pitch their ac-

tions as worthwhile "reforms." As a result, voters may end up experiencing real regret if they end up backing a "wolf in sheep's clothing." Likewise, in Luo & Przeworski's (2022) model backsliding also proceeds "by stealth," such that voters may not realize until it is too late that an incumbent who has broken norm after norm is firmly entrenched in power. Or, as Grillo & Prato's (2021) model shows, politicians may even be able to perversely exploit citizens' commitment to democracy, playing on their "fear and relief" response by threatening to break democratic norms and then partially back down in order to elicit a boost in popular support. The lesson across these contributions is that, even if citizens care deeply about protecting democracy, politicians can easily manipulate their uncertainty and thus undermine the enforcement of norms.

Yet a fourth type of hindrance to accountability emerges when citizens cannot agree amongst themselves about whether transgressions have occurred and/or which party is responsible. Building on Weingast's (1997) classic coordination model of democracy and the rule of law, Carey et al. (2019) posit that deterring leaders from norm-breaking behavior depends 1) on whether citizens agree not only on which norms are worth defending, and 2) but also on whether citizens agree about whether norms have, in fact, been violated (and by whom). Drawing on survey data collected by Bright Line Watch from the first two years of the Trump presidency, they conclude that Americans largely met the first condition, but not the second.

Finally, a robust literature on misperceptions shows both that citizens systematically under-estimate out-partisan support for democracy, but also that correcting such misperceptions substantially reduces respondents' own willingness to support political violence (Lees & Cikara 2020; Pasek et al. 2020; Ruggeri et al. 2021; Mernyk et al. 2022; Westwood et al. 2022, but also see Druckman 2023). Notably, in Stanford's "Strengthening Democracy Challenge" Megastudy, these sorts of correction interventions proved the single most effective treatment for reducing anti-democratic attitudes (Braley et al. 2023; Voelkel et al. 2022; Mernyk et al 2022). And, as we describe in the section on identifying norms below,

precisely because citizens' mutual beliefs about each other are experimentally manipulated, such studies are also among the first to causally establish the conditional nature underlying democratic norms at the mass level.

Politicians to Citizens

Not nearly as central to contemporary debates on democratic backsliding, an older tradition in political science effectively flips the vertical relationship between citizens and politicians on its head. From Pareto to Schumpeter to Lasswell to Lipset, this tradition presumes it is the elites who hold the strongest pro-democratic commitments (for overviews, see Borchert 2010; Volpe 2021). One version of the classic custodial argument is that, even if democracy and its constraining institutions were originally designed by elites for purely instrumental reasons, continued adherence to democracy eventually "enlightened" such elites relative to the "ignorant" populace.[3]

While such theories fell out of fashion by the late 20th century, the recent turn to populism has reignited concerns that many citizens genuinely desire undemocratic behaviors from their representatives, such as lying and norm-breaking (Hahl et al. 2018) or expedient rule-bending (Hibbing & Theiss-Morse 2002; Bustikova & Guasti 2019; Bloeser et al. 2022). Whether more democratically-minded elites can thus help to counter such tendencies has been raised as a possible solution. To give a well-known example, during the 2008 U.S. election town hall meeting a Republican supporter made a racist statement about Barack Obama. John McCain was widely lauded for pushing back, stating "[Obama is] a decent family man and citizen that I just happen to have disagreements with on fundamental issues and that's what this campaign is all about."[4] Yet, the evidence that politicians' appeals to the public's "better angels" work more generally appears somewhat mixed (for positive results see Rinscheid et al. 2021; Chong & Druckman 2007; for weaker or null results on

[3]See Stouffer (1955), Prothro & Grigg (1960) and Sullivan et al. (1993) for the view that elites are more supportive of civil liberties and tolerant of opposing views than citizens.

[4]Martin & Parnes, "McCain: Obama not an Arab, crowd boos", *Politico*, October 10, 2008, https://www.politico.com/story/2008/10/mccain-obama-not-an-arab-crowd-boos-014479

persuasion, see Kalla & Broockman 2018).

Still others have called for elites to provide a more institutional solution, arguing, for instance, that "responsible" parties should work to screen out popular "would-be" authoritarian candidates (Levitsky & Ziblatt 2018; Rosenbluth & Shapiro 2018). Yet, even if parties are able to narrow the field for voters, acting like an elite cartel may simply further erode public trust in democratic institutions (Rhodes-Purdy & Rosenblatt 2023). More fundamentally, for any form of custodialism to operate effectively, it cannot rely on the shaky premise that politicians will act consistently against their own self-interest.

Politicians to Politicians

Indeed, as much of the literature on the collapse of democratic norms among elites suggests, politicians' self-interest often —although not always —cuts in the wrong direction. Consider Levitsky & Ziblatt's (2018) recounting of the "fateful alliances" in which existing political parties accept a kind of Faustian bargain with would-be autocrats in order to remain electorally relevant. Such mistakes, they argue, have been all too common historically: They were committed in inter-war Italy and Germany, in Venezuela under the veteran politician Rafael Caldera, who paved the way for Hugo Chavez's rise to the presidency in the 1990s, as well as in the contemporary United States with the repeated failures of the Republican Party to break ties with Donald Trump.

Likewise, scholars have also pointed out that democratic norms governing behavior between parties suffer whenever it is no longer in politicians' self interest to abide by them. In their description of the breakdown of mutual toleration and forbearance between opposing parties in the United States, Levitsky & Ziblatt (2018) cite the Republican party's fears of eventually losing its majority due to demographic shifts as one of the key drivers of elite discord. Alternatively, Lee (2016) posits that the "logic of confrontation," is instead rooted in how competitive the Republican Party became at the end of the 20th century. In her view, it is precisely because Republicans steadily gained a chance at winning control of Congress

that cooperation with Democrats has become both less necessary and less attractive (also see Mann & Ornstein 2016).

To partly help adjudicate these claims, Helmke et al. (2021) develop a formal model that pins down the constellation of constitutional and electoral conditions under which parties abide by norms of forbearance. Their basic argument is that the sorting of rural voters into the Republican Party has activated latent asymmetries contained in the U.S. Constitution, which allows the Republican Party to trample norms without fear of symmetric retaliation from Democrats (also see Levitsky & Ziblatt 2023).

Beyond parties, other scholars have explored how independent courts can help to preserve or undermine democratic norms among political elites in other branches. Recent theoretical work by Howell et al. (2023), for example, paints an especially grim, if paradoxical, picture: precisely because judges in the U.S. have so much power to reign in the executive, they are often quite reluctant to use it. The result is that executive aggrandizement, and, hence, democratic backsliding, accrues over time — not despite the separation of powers, but because of it. More optimistically, Staton et al. (2022) suggest that independent courts can yet offer some protection against democratic backsliding via their ability to indirectly encourage norms of prudence (i.e., forbearance) among politicians. To make their point, they then engage in a counterfactual exercise of imagining the norms that might have been broken had the courts been less independent, particularly with respect to Trump's attempts to overturn the 2020 election. Of course, the more politicized the U.S. judiciary becomes, the less capacity it will have to play any role —direct or indirect—in preserving the norms that sustain democracy.

Citizens to Citizens

The fourth and final strand of the literature that we consider focuses instead on how citizens' desire for social conformity with their peers shapes their willingness to follow democratic norms. Cuing social pressure has been shown to induce compliance with basic pre-

scriptive democratic norms, such as turning out to vote (Gerber et al. 2008) and expressing support for democracy (Bischof et al. 2023). Along these lines, Goldstein's (2024) theory of democratic norms effectively reverses Kuran's (1987) seminal work on tipping points in dictatorships, arguing that if sufficient numbers of citizens value democracy, their more authoritarian peers may be pressured to falsify their anti-democratic preferences and support pro-democratic norms instead. This, of course, requires that a) a large enough proportion of citizens intrinsically value norms, b) that their more autocratic peers indeed care what the majority thinks, c) and/or that members of the majority are willing to sanction non-conformity in others.

The emerging literature also recognizes that social norms may just as easily work in the other direction. That is, information shocks from elections or from other political events may instead free people with latent authoritarian preferences to suddenly begin to express them. As we elaborate more below, Bursztyn et al. (2020), for example, show that respondents who receive information that their region voted for Trump in 2016 are no more likely to opt to donate to xenophobic organizations — but they are substantially less likely to worry if those donations are made public. In a similar vein, Dinas et al. (2024) develop a series of experiments, also elaborated on below, showing that respondents in neighborhoods displaying the Spanish nationalist flag after the referendum on Catalonian independence were significantly less likely to hide their own underlying preferences for authoritarianism.

But even in contexts where most people do widely support democratic norms and recognize anti-democratic behavior, they may be reluctant to actually enforce norms against peers who violate them. This is clearly the lesson that comes out of Alvarez-Benjumea & Valentim's (2023) novel experiment exposing subjects to different images of a citizen wearing political t-shirts, including those supporting the radical-right. Most respondents view wearing the radical-right t-shirt as a proverbial "finger in the eye" of the proscriptive norm to not publicly display support for proto-authoritarian political parties, but ultimately few were willing to say that they would do anything to actively censure such behavior.

Taken together, the upshot across these literatures is that democratic norms are fragile on multiple fronts. Whether transgressed by politicians or by peers, citizens are often either unwilling or unable to protect democracy. Likewise, elites can be counted on to push back against anti-democratic behavior only when it is compatible with their own self-interest to do so. What is more, challenging norm-breaking by one group of elites often necessitates that the other group break norms in turn. Indeed, once democratic norms are breached, attempts to enforce them *ex post* may only lead to further erosion and, as we discuss below, to the creation of new "post-democratic" norms in their place.

Identifying Norms

At root, identifying democratic norms entails establishing the extent to which individuals' willingness to support or jettison certain attitudes or behaviors varies according to their higher-order beliefs. Within any democracy, of course, motivations are likely to be mixed. Surely some citizens (and even, perhaps, the occasional politician) may defend or abandon democratic principles no matter what others do. Consider, for instance, models in which citizens' decisions depend purely on how important protecting democracy is relative to other considerations, such as party and policy outcomes (i.e., as in Svolik & Graham's (2020) canonical theory). Here, an individual's utility function is simply composed of her preferences over policy and/or party and, say, a valence (e.g., values-based) component representing her taste for democratic norms. Enforcement is modeled as a binary choice (ie. a citizen can punish a would-be autocrat ($a_i = 1$) or not ($a_i = 0$)), such that

$$u(a_i) = \alpha_1(a_i) + \alpha_2(a_i) + ... + \alpha_n(a_i) = \alpha * a_i \tag{1}$$

where $\alpha_j \in \mathbb{R}$ is a weight on a particular dimension; $\alpha > 0$ implies that a citizen enforces democratic behavior as the benefits outweight the costs.

Measuring then how much "value" respondents place on democratic principles is rela-

tively straightforward, at least in an experimental setting. Conjoint analysis, for example, offers a clear method that enables researchers to randomize a wide variety of candidate attributes to explore systematically how respondents make tradeoffs between democracy, party, and policy (Svolik & Graham 2020; Carey et al. 2022).[5]

To identify democratic norms as conditional preferences in the sense we have been describing, however, requires a different approach. That is, researchers must take additional steps to show that support for norms depends either on actors' beliefs about what other actors will do (i.e., empirical expectations), and on their beliefs about what others believe they ought to do (i.e., normative expectations).

Returning to Weingast's (1997) baseline model, citizens' ability to punish rulers who transgress the rule of law hinges entirely on which action they expect other citizens to take. If their expectations are such that citizens know that each other will punish transgressions, then leaders can be deterred from violating norms in the first place. If, instead, neither side believes that the other will punish, then the leader gets away with her transgression. Empirical expectations (i.e., first-order beliefs) are thus fundamental to any basic coordination situation, where one's payoffs depend on what one expects others will do.

The social norms literature, however, stresses another type of conditionality based on second-order beliefs, in which payoffs reflect the desire for conformity itself. As Bicchieri explains, insofar as social norms offer players an additional psychological payoff for matching the actions of others, normative expectations effectively transforms what was originally a "mixed-motive game" (i.e., a game in which cooperative outcomes are not equilibria, such as a prisoner's dilemma) into a pure coordination game (Bicchieri 2006:1-54).

Normative expectations map a context C (e.g., the circumstances at play, the history between actors, social statuses, etc), the expected actions of others a_{-i}, and a potential action a_i into a normative evaluation (Sugden 1989; Bicchieri 2006). Thus a norm N gives

[5] For a critique of current interpretations of conjoint experiments, see Abramson et al. (2022).

a payoff to actors based on the appropriateness of their action.

$$N : C \times a_{-i} \times a_i \to \mathbb{R} \tag{2}$$

From a modeling standpoint, this yields a more complex utility function consisting of unconditional preferences (as above) and both types of conditional preferences, where the e and n superscript denotes the payoffs from strategic situations involving empirical and normative considerations, respectively.:

$$U(a_i, a_{-i}) = u(a_i) + u^e(a_i, a_{-i}) + u^n(a_i, a_{-i}) \tag{3}$$

Substantively, we should emphasize that the force of normative expectations is quite flexible. It may capture external sanctions imposed by the other actors for not conforming (e.g., ostracism), or it may be a more internalized psychological disutility (shame, guilt, awkwardness, see Elster 1994) that stems from balking at the social consensus. Some formal models thus express norms as an internal desire to be regarded by others as the right type, say, the social benefit of i choosing some action a_i is $E(t_i|a_i)$, which is how others update about her type t_i (Bursztein, et al. 2020, Benabou & Tirole 2011; Goldstein 2024). Others (Acemoglu & Jackson 2015; Bueno de Mesquita & Shadmehr 2023) model it as a Keynesian beauty contest, where individuals wish to match a_i with the societal average action A_i (e.g., minimize $-(a_i - A_i)^2$); conformity "feels right".

From an empirical standpoint, the key challenge of identification lies in disentangling these various components. To this end, the growing literature on misperceptions offers one relatively straightforward formula for how researchers can begin to delineate the effects of empirical expectations by manipulating first-order beliefs. In a standard survey experiment format, respondents are 1) asked about out-partisan's views of democracy, 2) some portion of respondents are randomly treated with correct information about out-partisan views, and 3) a comparison is made between the mean responses of the treated and untreated respondents

in terms of their anti-democratic attitudes. Only if the average responses differ meaningfully across the two groups, can we infer that norms (i.e., at least those that revolve around information about first-order preferences) are operative. That is, we learn the extent to which an individual's willingness to sanction norm violations hinges on her expectations about what others would do. Conversely, if there is no difference between the control and treatment groups, descriptive norms are likely not at play.

Some studies go even further in this direction, including treatments that tap into normative expectations. Take, for instance, Bursztein et al. (2020) mentioned in the previous section, which studies anti-immigration preferences and social norms. Exploiting a natural experiment involving differing county-level versus metropolitan-level results from the 2016 presidential election, the authors randomly assign respondents to treatments in which they learn that either Clinton or Trump won the most votes. This serves as a signal for the popularity of anti-immigrant sentiments in their district. Next, they offer respondents the chance to donate a bonus dollar to an anti-immigrant organization. This is then followed by an additional cross-randomization treatment informing respondents that their donation will remain completely anonymous or not, which proxies for social image concerns. To establish whether normative expectations drive xenophobic behavior (versus, say, a change in individual preferences), the authors posit that the differences in donation rates between the Clinton and Trump groups will be isolated to public donations. Specifically, the gap between private and public donations will be much larger in the Clinton group, where the social stigma against supporting anti-immigrant causes remains in place.

A similar approach is adopted by Dinas et al. (2024). In this study, the authors use the density of Spanish flags across neighborhoods in Madrid as a signal about the social acceptability of publicly expressing support for nationalism and the Francoist dictatorship. To then ascertain the effect of this signal on the willingness of individual's to condone such norm-breaking behavior, they administer a list experiment to citizens from different neighborhoods to gauge their willingness to openly express anti-democratic sentiments. If

normative expectations are operative, the difference-in-difference between the direct question responses and the list responses should be greater in neighborhoods with fewer flags where the social stigma against supporting nationalism is presumably stronger and the incentives to engage in preference falsification are therefore heightened.

Taken together, the extant literature has made considerable headway in developing sophisticated methodological strategies to isolate the role of expectations; and, thus, in identifying causally the extent to which democratic norms are at play. Such contributions are impressive, exploiting deep empirical knowledge and deploying cutting-edge research designs. More substantively, the empirical findings reviewed above also suggests an interesting tension: shifting empirical expectations seems to actually change individual attitudes about anti-democratic norms (e.g., see Druckman 2023), while shifting normative expectations appears to merely free people to express their true attitudes. Whether this observation endures as a more robust regularity, or is purely an artifact of the particularities of the extant literature, remains an open question.

Finally, a few caveats about such methodological strategies are in order. First, most of these experiments rely on the rather strong assumption that a treatment modifies expectations (and nothing else). Furthermore, the experimental literature in economics also warns us that such manipulations are not entirely straightforward. Even in the lab, subjects often fail to correctly quantify the probability of simple events (Kahneman & Tversky 1974) due to a reliance on intuition and heuristic “rules of thumb.” Similarly, subjects may give misleading or wrong answers about their own behavior (Smallpage et al. 2023; Cloward 2014) or about the expected behavior of others (Aycinena et al. 2024, Bicchieri et al. 2023). A common recommendation is thus to monetarily incentivize answers, to both force higher cognitive effort as well as to reduce the incentives to misrepresent (Gachter & Renner 2010, Bicchieri & Xiao 2009). Still, if democratic norms “in the wild” ultimately operate more heuristically than rationally (cf. Bicchieri 2017), then the external validity of even the most cleverly designed experiments remains an issue.

Future Directions

The urgency of global democratic backsliding necessitates reflecting not just on why we should study democratic norms, but on how we should go about the task. Though norms have long been one of the central objects of study in the social sciences, contemporary political scientists have yet to fully converge on a definition of what norms are, or what makes them democratic. The first part of our essay seeks to remedy this by treating norms as conditional preferences rooted in shared expectations about what others will do, and shared beliefs about what others expect one ought to do. We are more agnostic when it comes to labeling, nudging scholars to draw on existing theories of democracy to thoughtfully assess how to best characterize a given norm. In addition, we have developed a basic typology to organize and review the burgeoning literature on the role of norms in democratic backsliding. Finally, using a simple general utility function, we have both laid out the problem of measuring democratic norms and reviewed multiple recent efforts to causally identify the role the play. The goal for the remainder of the essay is less reflective, than generative.

Going forward we see three main areas in which to build. First, to the extent that expectations are fundamental to identifying and understanding democratic norms, political scientists can do more to manipulate *whose* expectations matter. As cognitive scientists have long recognized, often the key to successful behavioral interventions lies in figuring out which peer or "reference group" is most salient (Miller & Prentice 1996; Tankard & Paluck 2016). From our perspective, how might shifting the identity of $-i$ affect the extent to which actors either alter or reveal their preferences over democratic norms? More generally, what happens when an individual is confronted with conflicting expectations — as in, their "friendly" network approves of flouting a pro-democratic norm, yet their "adversaries" do not, or vice versa?

To date, the democratic backsliding literature on misperceptions has largely revolved around correcting respondents' views of out-partisans, with the attendant theoretical focus

on mechanisms like threat and retaliation (Braley et al. 2022). An obvious candidate for future research thus lies in exploring the effects of updating views about co-partisans. For example, in their replication of Pasek et al. (2022), Bright Line Watch (2021) found that, on average, respondents tended to also substantially underestimate the value that their co-partisan placed on norms such as "using agencies to punish political rivals," "respect for free and fair elections," and "protecting equal, political and legal rights." Exploring whether correcting information about in-group values would have a similar – or perhaps an even greater – effect than correcting information about out-group preferences could thus help to disentangle more positive drivers of norms, such as the desire to mimic a congenial group.

A second area of future inquiry focuses on normalization. Whereas initial studies of democratic backsliding appropriately highlighted problems of uncertainty around interpreting the meaning of norm-breaking by outsider politicians (e.g., recall the 2016 debates about whether Trump's rhetoric meant to be taken "literally or just seriously"), the later stages of backsliding arguably present potential "enforcers" with a very different sort of challenge. Once individuals have sufficient information, how do they process it? Do they simply seek to reduce cognitive dissonance by re-framing violations, or by "elevating" other considerations, such as the well-being of the country (e.g., Krishnajaran 2022)? Or, are violations, particularly when they become incessant, instead normalized? Using a novel experiment in which subjects face repeated exposure to Trump's tweets impugning the integrity of the 2020 election, for example, Clayton et. al. (2022) explore whether this also alters perceptions of past norm violations under previous governments. Although they find no effects, future research could certainly build on their research design, while also operationalizing "normalization" in alternative ways.

Finally, by focusing primarily on the role of norms in the context of democratic backsliding we have necessarily privileged a particular phase of the so-called "life cycle" of political norms. That is, we have largely concentrated on accounts of how seemingly long-lived and well-established democratic norms suddenly come to die. A key task going forward is to ex-

pand our focus to include whatever new “post-democratic” norms are being created in their stead. Beyond the normalization process referenced above, norm creation often involves a more explicit attempt to shift beliefs not only about what is likely to happen in a given situation, but also about what sorts of behavior now warrant sanctioning. Studying how this occurs —how “entrepreneurial” individuals, whether they be elected politicians, bureaucrats, or other prominent members of civil society guide these processes, how quickly observers fall into line, and how push-back collapses—will be an important part of the unenviable future of studying post-democratic societies.

Acknowledgments

For comments on an earlier draft, we are extraordinarily grateful to Jamie Druckman, Hein Goemans, Susan Stokes, Brendan Nyhan and participants at the conference, “Democratic Frontsliding” at Princeton University, April 2024.

References

[1] Scott F Abramson, Korhan Koçak, and Asya Magazinnik. What do we learn about voter preferences from conjoint experiments? American Journal of Political Science, 66(4):1008–1020, 2022.

[2] Daron Acemoglu and Matthew O Jackson. History, expectations, and leadership in the evolution of social norms. The Review of Economic Studies, 82(2):423–456, 2015.

[3] Amel Ahmed. Is the american public really turning away from democracy? backsliding and the conceptual challenges of understanding public attitudes. Perspectives on Politics, 21(3):967–978, 2023.

[4] Gabriel Abraham Almond and Sidney Verba. The civic culture: Political attitudes and democracy in Five nations. Princeton University Press, 1963.

[5] Amalia Alvarez-Benjumea and Vicente Valentim. The enforcement of political norms. British Journal of Political Science, pages 1–24, 2023.

[6] Kenneth J Arrow. The organization of economic activity: issues pertinent to the choice of market versus nonmarket allocation. The analysis and evaluation of public expenditure: the PPB system, 1:59–73, 1969.

[7] Robert Axelrod. An evolutionary approach to norms. American Political Science Review, 80(4):1095–1111, 1986.

[8] Diego Aycinena, Francesco Bogliacino, and Erik O Kimbrough. Measuring norms: Assessing the threat of social desirability bias to the bicchieri and xiao elicitation method. Journal of Economic Behavior & Organization, 222:225–239, 2024.

[9] Julia R. Azari and Jennifer K. Smith. Unwritten rules: Informal institutions in established democracies. Perspectives on Politics, 10(1):37–55, Mar 2012.

[10] Larry M Bartels. Ethnic antagonism erodes republicans' commitment to democracy. Proceedings of the National Academy of Sciences, 117(37):22752–22759, 2020.

[11] Roland Benabou and Jean Tirole. Laws and norms. Technical report, National Bureau of Economic Research, 2011.

[12] Nancy Bermeo. On democratic backsliding. Journal of democracy, 27(1):5–19, 2016.

[13] Cristina Bicchieri. The grammar of society the nature and dynamics of social norms. Cambridge University Press, 2006.

[14] Cristina Bicchieri. Norms in the wild: How to diagnose, measure, and change social norms. Oxford University Press, 2016.

[15] Cristina Bicchieri, Eugen Dimant, and Silvia Sonderegger. It's not a lie if you believe the norm does not apply: conditional norm-following with strategic beliefs. Games and Economic Behavior, 2023.

[16] Cristina Bicchieri and Erte Xiao. Do the right thing: but only if others do so. Journal of Behavioral Decision Making, 22(2):191–208, 2009.

[17] Daniel Bischof, Tim Lars Allinger, Morgan Le Corre Juratic, and Kristian Vrede Skaaning Frederiksen. (mis-) perceiving support for democracy: The role of social norms for democracies. 2023.

[18] Andrew J Bloeser, Tarah Williams, Candaisy Crawford, and Brian M Harward. Are stealth democrats really committed to democracy? process preferences revisited. Perspectives on Politics, 22(1):116–130, 2022.

[19] Jens Borchert. "they ain't making elites like they used to": The never ending trouble with democratic elitism. In Heinrich Best and John Higley, editors, Democratic Elitism: New Theoretical and Comparative Perspectives, pages 23–41. Brill, 2010.

[20] P. Bourdieu and R. Nice. Distinction: Social critique of the judgement of taste. Routledge Kegan Paul, 1984.

[21] Samuel Bowles and Herbert Gintis. A cooperative species: Human reciprocity and its evolution. In A cooperative species. Princeton University Press, 2011.

[22] Robert Boyd and Peter J. Richerson. Origin and evolution of cultures (evolution and cognition). Oxford University Press, 2005.

[23] Alia Braley, Gabriel S. Lenz, Dhaval Adjodah, Hossein Rahnama, and Alex Pentland. Why voters who value democracy participate in democratic backsliding. Nature Human Behaviour, 7(8):1282–1293, May 2023.

[24] Geoffrey Brennan, Lina Eriksson, Robert E Goodin, and Nicholas Southwood. Explaining norms. Oxford University Press, USA, 2013.

[25] Dan Brinks. The rule of (non) law: prosecuting police killings in brazil and argentina.

In Gretchen Helmke and Steven Levitsky, editors, Informal institutions and democracy: Lessons from Latin America. Johns Hopkins University Press, 2006.

[26] Ethan Bueno de Mesquita and Mehdi Shadmehr. Social norms and social change. Quarterly Journal of Political Science, 18(3):339–363, 2023.

[27] Leonardo Bursztyn, Georgy Egorov, and Stefano Fiorin. From extreme to mainstream: The erosion of social norms. American Economic Review, 110(11):3522–48, November 2020.

[28] Lenka Buštíková and Petra Guasti. The state as a firm: Understanding the autocratic roots of technocratic populism. East European Politics and Societies, 33(2):302–330, 2019.

[29] John Carey, Katherine Clayton, Gretchen Helmke, Brendan Nyhan, Mitchell Sanders, and Susan Stokes. Who will defend democracy? evaluating tradeoffs in candidate support among partisan donors and voters. Journal of Elections, Public Opinion and Parties, 32(1):230–245, Jul 2022.

[30] John M Carey, Gretchen Helmke, Brendan Nyhan, Mitchell Sanders, and Susan Stokes. Searching for bright lines in the trump presidency. Perspectives on Politics, 17(3):699–718, 2019.

[31] Caterina Chiopris, Monika Nalepa, and Georg Vanberg. A wolf in sheep's clothes: Citizen uncertainty democratic backsliding. https://www.dropbox.com/scl/fi/vwi3d3z0esg2zt3mj38ul/DemocraticBacksliding.pdf?rlkey=7r71uwkoc25p07wtruc4syx2h&dl=0, 2023.

[32] Dennis Chong and James N Druckman. Framing public opinion in competitive democracies. American political science review, 101(4):637–655, 2007.

[33] Robert B Cialdini, Carl A Kallgren, and Raymond R Reno. A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In Advances in experimental social psychology, volume 24, pages 201–234. Elsevier, 1991.

[34] Katherine Clayton, Nicholas T. Davis, Brendan Nyhan, Ethan Porter, Timothy J. Ryan, and Thomas J. Wood. Elite rhetoric can undermine democratic norms. Proceedings of the National Academy of Sciences, 118(23), Jun 2021.

[35] Karisa Cloward. False commitments: Local misrepresentation and the international norms against female genital mutilation and early marriage. International Organization, 68(3):495–526, 2014.

[36] James S. Coleman. Foundations of Social Theory. Belknap Press of Harvard University Press, 1990.

[37] Robert A Dahl. Polyarchy: Participation and opposition. Yale university press, 1971.

[38] Elias Dinas, Sergi Martínez, and Vicente Valentim. Social norm change, political symbols, and expression of stigmatized preferences. The Journal of Politics, page 000–000, Mar 2024.

[39] James N. Druckman. Correcting misperceptions of the other political party does not robustly reduce support for undemocratic practices or partisan violence. Proceedings of the National Academy of Sciences, 120(37), Sep 2023.

[40] James N. Druckman, Suji Kang, James Chu, Michael N.Stagnaro, Jan G. Voelkel, Joseph S. Mernyk, Sophia L. Pink, Chrystal Redekopp, DavidG. Rand, and Robb Willer. Correcting misperceptions of out-partisans decreases american legislators' support for undemocratic practices. Proceedings of the National Academy of Sciences, 120(23), May 2023.

[41] Robert C. Ellickson. Law and economics discovers social norms. The Journal of Legal Studies, 27(S2):537–552, Jun 1998.

[42] Jon Elster. Social norms and economic theory. Journal of Economic Perspectives, 3(4):99–117, 1989.

[43] Jon Elster. Rationality, emotions, and social norms. Synthese, pages 21–49, 1994.

[44] Jean Ensminger and Jack Knight. Changing social norms: Common property, bridewealth, and clan exogamy. Current Anthropology, 38(1):1–24, Feb 1997.

[45] Clifford Geertz. The Interpretation of Cultures: Selected Essays. Basic Books, 1973.

[46] Michele J Gelfand, Sergey Gavrilets, and Nathan Nunn. Norm dynamics: Interdisciplinary perspectives on social norm emergence, persistence, and change. Annual Review of Psychology, 75(1):341–378, 2024.

[47] Alan S Gerber, Donald P Green, and Christopher W Larimer. Social pressure and voter turnout: Evidence from a large-scale field experiment. American political Science review, 102(1):33–48, 2008.

[48] Herbert Gintis. Social norms as choreography. politics, philosophy & economics, 9(3):251–264, 2010.

[49] Daniel A. N. Goldstein. The social foundations of democratic norms. http://dx.doi.org/10.2139/ssrn.4314977, October 2024.

[50] Matthew H. Graham and Milan W. Svolik. Democracy in america? partisanship, polarization, and the robustness of support for democracy in the united states. American Political Science Review, 114(2):392–409, Apr 2020.

[51] Edoardo Grillo and Carlo Prato. Reference points and democratic backsliding. American Journal of Political Science, 67(1):71–88, Nov 2021.

[52] Simon Gächter and Elke Renner. The effects of (incentivized) belief elicitation in public goods experiments. Experimental Economics, 13(3):364–377, Jul 2010.

[53] Oliver Hahl, Minjae Kim, and Ezra W. Zuckerman Sivan. The authentic appeal of the lying demagogue: Proclaiming the deeper truth about political illigitimacy. American Sociological Review, Jul 2018.

[54] Gretchen Helmke, Mary Kroeger, and Jack Paine. Democracy by deterrence: Norms, constitutions, and electoral tilting. American Journal of Political Science, 66(2):434–450, Oct 2021.

[55] Gretchen Helmke and Steven Levitsky. Informal institutions and comparative politics: A research agenda. Perspectives on Politics, 2(04):725–740, Dec 2004.

[56] Gretchen Helmke and Steven Levitsky. Introduction. In Gretchen Helmke and Steven Levitsky, editors, Informal institutions and democracy: Lessons from Latin America. Johns Hopkins University Press, 2006.

[57] John R Hibbing and Elizabeth Theiss-Morse. Stealth democracy: Americans' beliefs about how government should work. Cambridge University Press, 2002.

[58] Alisha C Holland. Forbearance as redistribution: The politics of informal welfare in Latin America. Cambridge University Press, 2017.

[59] Christine Horne and Stefanie Mollborn. Norms: An integrated framework. Annual Review of Sociology, 46(1):467–487, 2020.

[60] William G Howell, Kenneth A Shepsle, and Stephane Wolton. Executive absolutism: The dynamics of authority acquisition in a system of separated powers. Quarterly Journal of Political Science, 18:243–275, 2023.

[61] David Hume. A treatise of human nature. Oxford University Press, [1740] 2000.

[62] Aziz Huq and Tom Ginsburg. How to lose a constitutional democracy. University of Chicago Press, 2018.

[63] Joshua L Kalla and David E Broockman. The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. American Political Science Review, 112(1):148–166, 2018.

[64] Michihiro Kandori. Social norms and community enforcement. The Review of Economic Studies, 59(1):63–80, 1992.

[65] Jon Kingzette, James N Druckman, Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. How affective polarization undermines support for democratic norms. Public Opinion Quarterly, 85(2):663–677, Sep 2021.

[66] Suthan Krishnarajan. Rationalizing democracy: The perceptual bias and (un)democratic behavior. American Political Science Review, 2022.

[67] Timur Kuran. Preference falsification, policy continuity and collective conservatism. The Economic Journal, 97(387):642–665, 1987.

[68] Frances E Lee. Insecure majorities: Congress and the perpetual campaign. University of Chicago Press, 2016.

[69] Jeffrey Lees and Mina Cikara. Inaccurate group meta-perceptions drive negative outgroup attributions in competitive contexts. Nature human behaviour, 4(3):279–286, 2020.

[70] Steven Levitsky and Daniel Ziblatt. How democracies die. Broadway Books, 2019.

[71] Steven Levitsky and Daniel Ziblatt. Tyranny of the minority: Why American democracy reached the breaking point. Crown, 2023.

[72] David Lewis. Convention: A philosophical study. Harvard University Press, 1969.

[73] Seymour Martin Lipset. Some social requisites of democracy: Economic development and political legitimacy. American Political Science Review, 53(1):69–105, Mar 1959.

[74] Andrew T. Little and Anne Meng. Measuring democratic backsliding. PS: Political Science and Politics, page 1–13, Jan 2023.

[75] Zhaotian Luo and Adam Przeworski. Democracy and its vulnerabilities: Dynamics of democratic backsliding. Available at SSRN 3469373, 2022.

[76] Gerry Mackie. Ending footbinding and infibulation: A convention account. American sociological review, pages 999–1017, 1996.

[77] Thomas E Mann and Norman J Ornstein. It's even worse than it looks: How the American constitutional system collided with the new politics of extremism. Basic Books, 2016.

[78] Joseph S. Mernyk, Sophia Lerner Pink, James Druckman, and Robb Willer. Correcting inaccurate metaperceptions reduces americans' support for partisan violence. Proceedings of the National Academy of Sciences, 119(16), Apr 2022.

[79] Suzanne Mettler and Robert C. Lieberman. Four threats: The recurring crises of American Democracy. St. Martin's Griffin, 2020.

[80] Robert Mickey. Paths Out of Dixie: The Democratization of Authoritarian Enclaves in America's Deep South, 1944-1972. Princeton University Press, 2015.

[81] Dale T. Miller and Deborah A. Prentice. The construction of norms and social standards. In E. Tory Higgins and Arie W.Editors Kruglanski, editors, Social Psychology: Handbook of Basic Principles. Guilford Press, 1996.

[82] David Montgomery. The abnormal presidency. The Washington Post, Nov 2020.

[83] Douglass C North. Institutions. Journal of Economic Perspectives, 5(1):97–112, Feb 1991.

[84] Guillermo A O'Donnell. Illusions about consolidation. Journal of democracy, 7(2):34–51, 1996.

[85] Talcott Parsons, Edward Shils, and Edward Chace Tolman. Toward a general theory of action / Edward C. Tolman and others; Talcott Parsons, Edward A. Shils editors. Harper Row, 1951.

[86] Michael H. Pasek, Lee-Or Ankori-Karlinsky, Alex Levy-Vene, and Samantha L. Moore-Berg. Misperceptions about out-partisans' democratic values may erode democracy. Scientific Reports, 12(1), Sep 2022.

[87] Eric A. Posner. Law and social norms: The case of tax compliance. Virginia Law Review, 86(8):1781, Nov 2000.

[88] James W Prothro and Charles M Grigg. Fundamental principles of democracy: Bases of agreement and disagreement. The Journal of politics, 22(2):276–294, 1960.

[89] Adam Przeworski. Who decides what is democratic? Journal of Democracy, 35(3):5–16, 2024.

[90] Adam Przeworski, Michael Alvarez, José Antonio Cheibub, and Fernando Limongi. Democracy and development: Political institutions and well-being in the world, 1950-1990. Number 3. Cambridge University Press, 2000.

[91] Robert D. (1941) Putnam, Robert Leonardi, and Raffaella Y. Nanetti. Making democracy work: Civic traditions in modern Italy. Princeton University Press, 1994.

[92] Matthew Rhodes-Purdy and Fernando Rosenblatt. Raising the red flag: democratic elitism and the protests in chile. Perspectives on Politics, 21(1):241–253, 2023.

[93] Adrian Rinscheid, Silvia Pianta, and Elke U Weber. What shapes public support for climate change mitigation policies? the role of descriptive social norms and elite cues. Behavioural Public Policy, 5(4):503–527, 2021.

[94] Frances Rosenbluth and Ian Shapiro. Responsible parties: Saving democracy from itself. Yale University Press, 2018.

[95] Kai Ruggeri, Bojana Većkalov, Lana Bojanić, Thomas L Andersen, Sarah Ashcroft-Jones, Nélida Ayacaxli, Paula Barea-Arroyo, Mari Louise Berge, Ludvig D Bjørndal, Aslı Bursalıoğlu, et al. The general fault in our fault lines. Nature Human Behaviour, 5(10):1369–1380, 2021.

[96] Thomas C. Schelling. The Strategy of Conflict. Harvard University Press, 1960.

[97] Brian Skyrms. The stag hunt and the evolution of social structure. Cambridge University Press, 2004.

[98] Steven M Smallpage, Adam M Enders, Hugo Drochon, and Joseph E Uscinski. The impact of social desirability bias on conspiracy belief measurement across cultures. Political Science Research and Methods, 11(3):555–569, 2023.

[99] Chandra Sripada and Stephen Stich. A framework for the psychology of norms. In Peter Carruthers, Stephen Laurence, and Stephen Stich, editors, Innateness and the Structure of the Mind. Oxford University Press, 2007.

[100] Jeffrey K Staton, Christopher Reenock, and Jordan Holsinger. Can courts be bulwarks of democracy?: Judges and the politics of prudence. Cambridge University Press, 2022.

[101] Susan Stokes. Do informal rules make democracy work? accounting for accountability in argentina. In Gretchen Helmke and Steven Levitsky, editors, Informal institutions and democracy: Lessons from Latin America. Johns Hopkins University Press, 2006.

[102] Samuel Andrew Stouffer. Communism, conformity, and civil liberties: A cross-section of the nation speaks its mind. Transaction Publishers, 1955.

[103] Robert Sugden. Spontaneous order. Journal of Economic Perspectives, 3(4):85–97, Nov 1989.

[104] John L Sullivan, Pat Walsh, Michal Shamir, David G Barnum, and James L Gibson. Why politicians are more tolerant: Selective recruitment and socialization among political elites in britain, israel, new zealand and the united states. British Journal of Political Science, 23(1):51–76, 1993.

[105] Cass R. Sunstein. Social norms and social roles. Columbia Law Review, 96(4):903, May 1996.

[106] Margaret E. Tankard and Elizabeth Levy Paluck. Norm perception as a vehicle for social change. Social Issues and Policy Review, 10(1):181–211, Jan 2016.

[107] Alexis de Tocqueville. Democracy in America. part the second, the social influence of democracy. J. H.G. Langley, 1840.

[108] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. science, 185(4157):1124–1131, 1974.

[109] Edna Ullmann-Margalit. The emergence of norms. Clarendon Press, 1977.

[110] Vicente Valentim. Parliamentary representation and the normalization of radical right support. Comparative Political Studies, 54(14):2475–2511, Mar 2021.

[111] Ozan O Varol. Stealth authoritarianism. Iowa L. Rev., 100:1673, 2014.

[112] Jan G. Voelkel, Mashail Malik, Chrystal Redekopp, and Robb Willer. Changing americans' attitudes about immigration: Using moral framing to bolster factual arguments. The ANNALS of the American Academy of Political and Social Science, 700(1):73–85, Mar 2022.

[113] Jan G Voelkel, Michael Stagnaro, James Chu, Sophia Pink, Joseph Mernyk, C Redekopp, M Cashman, Q Submitters, J Druckman, D Rand, et al. Megastudy identifying

successful interventions to strengthen americans' democratic attitudes. Northwestern University: Evanston, IL, USA, 2022.

[114] Giorgio Volpe. Italian elitism and the reshaping of democracy in the United States. Routledge, 2021.

[115] Bright Line Watch. Democratic transgressions and constitutional hardball. http://brightlinewatch.org/democratic-transgressions-and-constitutional-hardball-bright-line-watch-october-20 //, October 2019.

[116] Bright Line Watch. A democratic stress test — the 2020 election and its aftermath. http://brightlinewatch.org/a-democratic-stress-test-the-2020-election-and-its-aftermathbright-line-watch-nove /, November 2020.

[117] Bright Line Watch. Tempered expectations and hardened divisions a year into the biden presidency. http://brightlinewatch.org/tempered-expectations-and-hardened-divisions-a-year-into-the-biden-presidency/, November 2021.

[118] Barry R. Weingast. The political foundations of democracy and the rule of the law. American Political Science Review, 91(2):245–263, Jun 1997.

[119] Sean J Westwood, Justin Grimmer, Matthew Tyler, and Clayton Nall. Current research overstates american support for political violence. Proceedings of the National Academy of Sciences, 119(12):e2116870119, 2022.

[120] H Peyton Young. The evolution of social norms. Annual Review of Economics, 7(1):359–387, 2015.