The Polarized Mind in Context: Interdisciplinary Approaches to the Psychology of Political Polarization
Jeroen M. van Baar and Oriel FeldmanHall

The Polarized Mind in Context: Interdisciplinary Approaches to the Psychology of Political Polarization

Jeroen M. van Baar¹ and Oriel FeldmanHall¹, ²

¹ Department of Cognitive, Linguistic, and Psychological Sciences, Brown University
² Carney Institute for Brain Science, Brown University

Existing research into the psychological roots of political polarization centers around two main approaches: one studying cognitive traits that predict susceptibility to holding polarized beliefs and one studying contextual influences that spread and reinforce polarized attitudes. Although both accounts have made valuable progress, political polarization is neither a purely cognitive trait nor a contextual issue. We argue that a new approach aiming to uncover interactions between cognition and context will be fruitful for understanding how polarization arises. Furthermore, recent developments in neuroimaging methods can overcome long-standing issues of measurement and ecological validity to critically help identify in which psychological processing steps—e.g., attention, semantic understanding, emotion—polarization takes hold. This interdisciplinary research agenda can thereby provide new avenues for interventions against the political polarization that plagues democracies around the world.

Public Significance Statement

To understand political polarization, it is key to study how people’s cognitive traits, such as intolerance to uncertainty, interact with contextual factors, such as social media. New tools from neuroscience can reveal how cognitive-contextual interactions shape the psychological processing of political information at multiple distinct processing steps, including attention and emotion. This interdisciplinary approach illuminates the psychological roots of polarization.

Keywords: political polarization, cognitive-contextual interactions, interdisciplinary approach, neuroimaging

States, electoral statistics expose several key polarization trends. This includes the increasing alignment of within-party opinions on distinct issues (e.g., same-sex marriage and gun control; Baldassarri & Gelman, 2008; Hetherington, 2001, 2009; Levendusky, 2009); a widening ideological gap between opposing parties associated with stronger party brands (Americans for Democratic Action, 2019; Fiorina et al., 2011; Poole & Rosenthal, 1984, 1991; Stonecash et al., 2003); and the homogenization of voting districts along lines of race, education, socioeconomic status, and other demographic factors (Frey, 1979; Grodzins, 1957; Issacharoff & Nagler, 2007; McLean, 2015; Stonecash et al., 2003; Tam Cho et al., 2013). These trends have sparked new research on the roots of polarization and extremist beliefs (e.g., McCauley & Moskalenko, 2017; Oosterhoff et al., 2018).

Although the field of psychology is situated to provide a deeper understanding of these trends by illuminating how an individual adopts polarized attitudes, our psychological grasp of political polarization remains relatively limited.
This is partly because the psychological processes that shape polarization are difficult to reproduce in static, antiseptic laboratory settings. In addition, key psychological processes such as emotional responses to political information (Neumann et al., 2007) can be difficult to detect with classic measurements like verbal report (Nisbett & Wilson, 1977). This is particularly true for naturalistic socioemotional experiences that unfold rapidly, such as watching television (Cheong et al., 2020), and for the modern experience of political discourse via the Internet (McDuff et al., 2013). Given these constraints, it is unclear in which psychological processing steps—for example, attention, perception, language understanding, emotion, decision-making—polarization takes hold.

To date, psychological accounts of polarization broadly fall into two frameworks that each focus on a specific, readily observable aspect of polarization. The first account examines how individual trait differences, such as need for closure and cognitive inflexibility, are associated with holding polarized attitudes (Jost et al., 2003; Rollwage et al., 2019; Zmigrod, 2020). The second account seeks to understand how contextual factors such as media filter bubbles and biased social networks shape polarized political behavior (Iyengar & Ansolabehere, 1995; Kreps, 2020; Pariser, 2011; Starbird et al., 2014; Stroud, 2010). While both accounts have made valuable progress in describing cognitive and contextual concomitants of polarization, political polarization is neither a purely cognitive trait nor a contextual issue.

Instead, it is more likely that polarization arises from the complex interplay of cognition and context (Hatemi & McDermott, 2016; Jung et al., 2019; Zaller, 1992), which together influence specific psychological processes in the cognitive hierarchy. Adopting a framework that interrogates these interactive effects of context and cognition, combined with methods that can measure specific processing steps as they unfold in real time, can reveal new insight into the psychological mechanisms of political polarization. Advanced neuroscientific methods can help overcome long-standing issues of measurement and ecological validity in this endeavor. As just one example, using cutting-edge neuroimaging approaches, we can now measure the response of distinct psychological processes in the minds of participants undergoing relatively naturalistic political experiences in the lab (e.g., watching television), which sidesteps the need to rely on static experimental stimuli and biased self-report measures. Such methods give new freedom to exploring the landscape of polarization and can help us better understand how political polarization has become such a large-scale societal phenomenon.

Individual Cognitive Traits Contribute to Political Polarization

Leading cognitive-based theories in political psychology argue that a person’s idiosyncratic psychological needs and traits contribute to the adoption of particular political views (Jost et al., 2003; McDermott & Hatemi, 2017). A now classic theory reasons that a conservative worldview can satisfy the “epistemic need” for a predictable, organized social environment and clear-cut principles about how the world works (Baron & Jost, 2019; Hibbing et al., 2015; Jost & Amodio, 2012; Jost et al., 2009; Jost et al., 2003; Jost et al., 2007). For instance, conservatives are thought to have a greater need for closure (Baron & Jost, 2019; Jost et al., 2003; Jost et al., 2017), which breeds an overall aversion to change. Recent work, however, suggests that strong epistemic motivation and lack of cognitive flexibility are found not just in conservatives but in anyone on the extreme end of the political spectrum—committed liberals and conservatives alike (Ditto et al., 2019; Rollwage et al., 2018; Rollwage et al., 2019; van Prooijen & Krouwel, 2019; Zmigrod, 2020; Zmigrod et al., 2019). These epistemic needs may be compounded by social desires such as the need to belong (Baumeister & Leary, 1995; Kunst et al., 2019), which strengthens the motivation to hold beliefs that maintain a good position within a desired social group (Correll & Park, 2005; Tetlock, 2002; Van Bavel & Pereira, 2018).

How do the psychological needs stemming from our cognitive traits exacerbate polarization? First, if a view is held in order to satisfy a need, that view becomes a valued possession (Abelson, 1986) and contradictory views become threatening. For instance, a person with low cognitive flexibility may value the belief that their party is always right and will subsequently reject evidence of corruption among
their party’s leaders. Second, psychological needs can strengthen existing beliefs through motivated reasoning (Kunda, 1990). People can preferentially access information from memory that supports a desired view of the world (Bower, 1981; Festinger & Carlsmith, 1959; Kunda, 1987), seek out external evidence that uniquely supports desired beliefs (Bakshy et al., 2015; Caddick & Rottman, 2019; Campbell & Kay, 2014; De Dreu et al., 2008; Frimer et al., 2017; Garrett, 2009a, 2009b; Gilovich, 1983; Lord et al., 1979; Nyhan & Reifler, 2010, 2015; Sharot et al., 2011; Stanley et al., 2019; Wood & Porter, 2019), and apply specific rules to make judgments or choices that support desired goals (e.g., ignoring the base rate; Ginossar & Trope, 1987). Third, once a person has developed a biased worldview, any new information is interpreted in a partisan way even without motivated reasoning (Cook & Lewandowsky, 2016; Gerber & Green, 1999; Jern et al., 2014; Miller & Ross, 1975). This is because unbiased (or “rational”) belief updating depends on an individual’s priors, which are constructed on what has been learned about the world in the past (Gerber & Green, 1999). Because diverging priors about the causal structure of the world can drive belief polarization, it is crucial to understand how an individual’s social and political context can generate biased priors.

**Contextual Factors Shape Political Polarization**

Context is also known to be a powerful determinant of polarization (Bail et al., 2018; Brady et al., 2017; Johnson et al., 2019; Pomerantsev, 2019; Stroud, 2010; Urman, 2019). Historically, examining contextual influences on behavior has been met with some controversy as the power of the situation undermines the common assumption in experimental psychology that people exhibit uniform and coherent behaviors (Fiske, 2018; Ross & Nisbett, 1991). However, in the past few decades, it has become well established that personality traits have relatively low predictive value for real-world behavior when viewed in isolation, even when studying how behavior generalizes between well-controlled laboratory tasks such as economic games (Galizzi & Navarro-Martinez, 2018; Pedroni et al., 2017). Context thus remains a critical factor for understanding social-psychological phenomena, such as political polarization (Jost et al., 2009). We refer to context as the structure of the social environment, which includes factors known to influence behavior, such as social pressure, others’ expectations, social norms, and habits (Eagly & Chaiken, 1993). These factors are exogenous to the individual and can be as low-level as a nuclear family’s media consumption habits or as high-level as the diversity of one’s social network. This definition of context is deliberately broad, since many types of contextual influences are critical in producing polarization. The key unifying aspect is that context provides information about the political world that can instrumentally shape polarized attitudes.

For example, partisans’ selective exposure to biased news and opinions is a key contextual factor driving polarization (Lord et al., 1979; Stroud, 2010; Vallone et al., 1985; Zaller, 1992). In a matter of decades, the ways in which Americans get their news has changed dramatically. More than ever before, social media leads as the number one source of news (Barthel, 2019; Shearer, 2018; Shearer & Matsa, 2018). The abundance of online news sites has moved the selection and sorting of information from the newsroom editorial board to citizens themselves and to algorithms created by social media platforms (Bakshy et al., 2015; Bennett & Iyengar, 2008). These algorithms are designed to pick posts that the user will like, which means social media users are becoming increasingly exposed to views they already agree with. This is known as the “filter bubble” effect (Bakshy et al., 2015; Bennett & Iyengar, 2008; Bovet & Makse, 2019; Flaxman et al., 2016; Huckfeldt et al., 2004; Pariser, 2011; Stewart et al., 2019; Stroud, 2010). The online spread of extreme beliefs within ideological groups can be further hastened by including moral-emotional statements in social media posts (Brady et al., 2017; Heath et al., 2007). The more morally and emotionally salient a political tweet, the more likely it is to be shared among peers who already agree with the content, effectively producing a politically charged echo chamber.

Although less obvious than a social media echo chamber, the underlying network structure of our online and offline social environments can also bias our perceptions of politics (Johnson et al., 2019; Stewart et al., 2019). The relational structure of a community—who communicates with whom
—can yield an uneven spreading of information about what other people in the community believe (Banerjee et al., 2013). In politics, this phenomenon can lead to information gerrymandering (Stewart et al., 2019), whereby voters adjust their voting behavior depending on how they believe others will vote. To make it concrete, suppose voters from two parties would prefer their own party to win an election, but would also prefer the other party to win over a situation of gridlock due to a split parliament. In this case, if one can convince voters of one party that the opposing party will most likely win, this would motivate people to vote for the opposing party out of a “realist” view that gridlock can be avoided (Stewart et al., 2019). This is not just hypothetical, as voter turnout was a decisive factor in the 2016 U.S. presidential election (Cohn, 2016). Similarly, the structure of a social network can also polarize beliefs over time: A low number of social connections can cause people to adopt polarized beliefs, while a high number of connections buffers against extreme views through repeated exposure to diverse attitudes (Grim et al., 2012). Finally, network structure plays a key role in determining the onset and extent of political activity such as intergroup conflict (Glowacki et al., 2016). Simply put, even the network structure of our social environment can have a profound impact on shaping our political opinions and actions.

Limitations of Existing Approaches

Despite great progress in our understanding of the cognitive and contextual bases of polarization, their explanatory power is limited when they are viewed in a siloed manner. It would be difficult, for example, to believe that the increase in political polarization over the last several decades was caused by a wholesale uptick in cognitive inflexibility across the population. Similarly, billions of people make use of social media, but not everyone adopts polarized political views. Polarization interventions also provide no evidence that examining context and cognition in isolation produces the level of understanding needed to actively mitigate polarization. Based on extant research about polarizing environments, one would predict that publishing corrections alongside biased information on the Internet—as is increasingly common on Twitter and Facebook—might reduce polarization. However, researchers have found that such interventions often fail to reduce misperceptions among the targeted ideological group, sometimes even backfiring to increase polarization (Bail et al., 2018; Jones & Harris, 1967; Nyhan & Reifler, 2010; but see Wood & Porter, 2019). Similarly, banning hate speech on social media like Facebook can make matters worse by pushing hate groups into global “dark pools” where hate-mongering messages can flourish unpolicied (Johnson et al., 2019). Other researchers have argued that cognitive training—for example, teaching individuals how to correctly interpret quantitative information about political issues—might reduce political polarization (Rollwage et al., 2019). Yet people with greater numeracy skills become more polarized when exposed to quantitative information about a political issue, suggesting an even greater capacity for conjuring up new, creative accounts of political events that support preexisting loyalties (Gaines et al., 2007; Kahan et al., 2017; Kahan et al., 2012). Together, these counterintuitive findings suggest there is a need for a new approach in examining political polarization.

The Framework: Cognition and Context Interact to Shape Political Polarization

An alternative possibility is that cognition-context interactions powerfully contribute to political polarization. This echoes an age-old debate in medicine on the impact of traits versus environment in disease etiology, colloquially termed nature versus nurture. For example, we know that certain genes can raise or lower our risk of developing cardiovascular disease. However, we also know that smoking is linked to cardiovascular disease. More recently, doctors have learned that these two risk factors interact in ways that can determine which medical intervention will be most successful. A gene called APOE can increase the risk of developing cardiovascular disease—but this is only true for people who smoke. Nonsmoking APOE carriers run the same risk as everyone else in the population (Talmud, 2007). Trait-environment interactions are found in clinical psychology as well, where they can have particularly counterintuitive effects. For instance, the orchid-dandelion hypothesis argues that children thought to be the most genetically vulnerable to psychological stressors (i.e., the orchids) actually show better mental health outcomes than their peers (i.e., the dandelions) if given the right support and protection during childhood (Boyce & Ellis, 2005; Dick et al., 2011; Lionetti et al., 2018).

Given the potential for interactions between cognitive and contextual drivers of polarization, it is likely that the success of stymieing political polarization will also depend on interventions that speak to trait-environment interactions (Jost et al., 2009; McDermott et al., 2013; Zaller, 1992). For instance, epistemic needs like intolerance of uncertainty may only enable motivated reasoning if one’s environment offers a broad enough array of political information. To make this idea concrete, consider the polarizing effect of “fake news” (Lazer et al., 2018). By inundating citizens with conflicting information sources (real or not), individuals who carry cognitive “risk factors” for polarization (e.g., high need for closure) are provided with the ambiguous information required to selectively bolster beliefs that support their partisanship (McDermott, 2019; Pomerantsev, 2019). Similarly, the epistemic desire to engage in motivated reasoning is amplified by the social need to justify one’s
beliefs to a homogeneous social environment where everyone believes the same thing (Tetlock, 1992, 2002). This can accelerate the spread of motivated beliefs, especially in tight-knit, ideologically homogeneous social networks. Moreover, after years of one-sided news consumption, the infrequent exposure to alternative views is increasingly met with a negative affective response (McDuff et al., 2013; Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). This bolsters the psychological need for the consumption of confirmatory evidence, completing a vicious cycle of cognitive and contextual drivers of polarization. These examples of interactions between cognitive predispositions and contextual triggers of polarization are reminiscent of gene-environment correlations (rGE) in behavioral genetics (Jaffee & Price, 2007). In rGE, one’s genotype partially determines the environmental influences to which one is exposed, such as when people with a genetic predisposition to extraversion seek out more diverse social environments. We suggest that similar interactive effects may powerfully drive polarization as cognitive traits determine exposure to, and processing of, inflammatory political information from the environment.

We propose a two-pronged research agenda that puts interactions between cognitive and contextual factors center stage. Cognitive-contextual political psychology (see Figure 1) aims to understand how cognition and context interact to yield political polarization. How do partisan cognitive traits and epistemic needs bias the processing of political information from the environment, and how does context shape the way we deploy relevant cognitive processes? Focusing on finding the answer to these questions can open up new avenues for research on political polarization. These avenues will benefit from multidisciplinary methods that can more effectively capture cognitive style (e.g., epistemic need surveys combined with tools from computational political psychology; Jost, 2017; Rollwage et al., 2019), polarized social networks (e.g., large-scale data sets scraped from Twitter or Facebook; Brady et al., 2017), and group dynamics (e.g., experimental tasks capable of measuring ingroup/outgroup influences; Stewart et al., 2019).

The Cognitive-Contextual Approach Provides a Deeper Understanding of Polarization

In 1954, Albert Hastorf and Hadley Cantril wrote a seminal paper on what they refer to as “a ‘real life’ study of a perceptual problem” (Hastorf & Cantril, 1954, p. 130). In describing spectators’ experiences during a particularly rough Dartmouth-Princeton football game, they reported that “Princeton students saw the Dartmouth team make over twice as many infractions as their own team made” (p. 130). Dartmouth students, on the other hand, saw both teams make about the same number of infractions and judged their own infractions to be rather mild. The researchers concluded, “It seems clear that the ‘game’ actually was many different games and that each version of the events that transpired was just as ‘real’ to a particular person as other versions were to other people” (p. 132). In other words, the students’ experience of the game was polarized, similar to how experiences of political events may be polarized between opposing partisans (Vallone et al., 1985; Zaller, 1992). Hastorf and Cantril hinted that such polarization was due to differences in the values, beliefs, and attentional biases that the spectators brought with them to the game, as well as the context in which they experienced football. For example, membership of a school and exposure to biased reporting in the college newsletter likely caused the students to bring different sets of prior knowledge and values to the game. On top of this, individual differences interacted with these priors to construe different events out of the occurrences on the field, which triggered a divergence in their subsequent psychological responses.

Translating this relatively low-stakes example to the present-day political realm, we see that voters increasingly

Figure 1

Psychological Approaches to Political Polarization

Note. Cognitive-contextual political psychology aims to reveal the interactions of cognitive traits with contextual influences that can drive polarized beliefs. See the online article for the color version of this figure.
consume a diet of biased news and social media posts. This is the contextual side of the coin. On the other hand, our cognitive architecture makes meaning out of information by integrating it with prior knowledge, which can lead to biased perceptions even if the same information is consumed. As we have seen, this cognitive integration process is thought to be exaggerated in individuals with low cognitive flexibility and strong epistemic needs, which feeds polarization in these specific individuals. Thus, cognitive-contextual interactions are paramount in explaining how the subjective experience of a sports game or a political event can be so different between opposing partisans. This biased subjective experience cannot be solely measured as a cognitive trait nor as a readily accessible feature of our environment. Instead, it takes shape inside our minds, where cognitive traits meet contextual input.

Using a cognitive-contextual lens to examine the political psychology literature can offer new insight into how political polarization arises. For example, a homogeneous political environment combined with a strong need to belong can yield simplistic, negative attitudes about the political outgroup, known as affective polarization (Iyengar et al., 2019; Iyengar et al., 2012). We experienced the outcome of cognition and context interacting when running a political experiment in our own lab. One strongly conservative participant, who had grown up in a homogeneous political environment and scored high on trait intolerance to uncertainty, emphatically explained that he was no longer interested in policy making—but rather in putting up a fight against the liberals who are out to “destroy his way of life.” More generally, political adversaries tend to ascribe outgroup aggression to hate, while ascribing aggressive ingroup behavior to love—a motive attribution asymmetry bias (Slovic et al., 2020; Waytz et al., 2014). These interpretations cannot both be true at once, demonstrating that our perceptions of others’ intentions are often inaccurate and instead align with how we want to see the world (Yudkin et al., 2019). Motive attribution asymmetry most likely arises from interactions between context (e.g., limited and negatively valenced exposure to the political outgroup) and cognition (for example, the epistemic motivation to see one’s ingroup as loving and one’s outgroup as hateful; Waytz et al., 2014). Such misconceptions generalize to beliefs about others’ judgments about the self (“group metaperceptions”): Political partisans tend to think that their adversaries feel more negatively about them than they actually do (Lees & Cikara, 2019; Vallone et al., 1985). This bias can lead to an overestimation of intentional obstructionism, which cyclically leads to the assumption that our opponents are acting in bad faith. These are prime examples of how a polarized perspective on the world is construed by cognitive processes that rigidly interpret new information through the biased lens of prior knowledge gleaned from a partisan context.

Polarized misperceptions extend beyond perceptions of self and other, which cause deleterious effects on democratic cooperation. For instance, war casualties (Gaines et al., 2007) and economic indicators (Bartels, 2002; Bullock et al., 2015) are perceived as good or bad news depending on one’s party affiliation. Even beliefs about objective facts differ between opposing partisans, revealing a “polarization of reality” (Alesina et al., 2020; Makridis & Rothwell, 2020). For instance, Republicans believe that the top 1% of wealth holders hold 53% of the wealth in the United States, while Democrats believe they hold 68% (the true number is around 42%; Stantcheva, 2020). These misperceptions, again, result from an interaction between contextual and cognitive influences as they can be altered by exposure to factual information, but participants are less motivated to consume such factual information if it challenges their views (Alesina et al., 2020). Since policy making relies on agreement about facts, the polarization of reality is detrimental to effective governance. The cognitive-contextual approach reveals that the polarization of our subjective experience of the political world—a phenomenon that is driven by the interaction of cognitive and contextual factors—is the Lynch pin of political polarization.

Political Polarization at Distinct Psychological Processing Steps

If cognitive and contextual influences interact to drive polarization, then a natural follow-up question is this: How does the interaction of cognition and context influence each discrete step in the psychological processing of political information? This mechanistic question requires drilling down into the distinct processing steps that sequentially turn political information into polarized attitudes (see Figure 2). We give three examples of how polarization may take hold in these processing steps: attention, semantics, and emotion. Other processing steps that may be implicated include (but are not limited to) perception, memory encoding, and decision-making.

First, polarization may arise at the level of attention, which is known to be driven by beliefs about which information will be most meaningful (Henderson & Hayes, 2018). Initial evidence using eye tracking reveals that motivated reasoning can affect political information search (Frimer et al., 2017), causing committed partisans to preferentially look at political posters depicting candidates and views from their own side (Marquart et al., 2016; Schmuck et al., 2019). It remains to be tested whether this form of polarized attention generalizes to more naturalistic political experiences such as watching news items or political debates on TV. One way to examine the interaction between cognition and context is to probe whether polarized attention is exacerbated by cognitive traits—including strong
epistemic needs that might motivate the individual to seek out confirmatory evidence (Frimer et al., 2017)—and a homogeneous social and media context. Together, these two factors should determine which pieces of information are meaningful and are more readily attended to (Henderson & Hayes, 2018).

Second, polarization may arise at the level of semantic language understanding. Language understanding heavily depends on contextual expectations, which drive how ambiguous words and sentences are interpreted (Hagoort & Indefrey, 2014; Huettig, 2015). For example, the word “jam” triggers an entirely different mental image when said in a music studio than when said in a breakfast room. This explains why the framing of political issues (which are highly complex and often morally ambiguous) can powerfully impact public opinion, such as when conservatives speak of a “death tax” instead of an “estate tax” or when liberals frame abortion as “a woman’s right to choose” (Lakoff, 2002; Scheufele, 2000). Contextual language expectations can be conceptualized as the coactivations of semantic nodes in a network of linguistic knowledge (McClelland & Rogers, 2003), where connection strengths can vary between individuals—for example, as a function of political affiliation (Halpern & Rodriguez, 2018). Since some political language is characterized as being highly ambiguous—politicians even use “dog-whistle” phrases that are intentionally vague yet carry a particular meaning to a subgroup of voters (Albertson, 2015; Haney López, 2015)—the way in which political language is semantically represented may vary strongly between opposing groups of partisans. If this is the case, semantic polarization may result from an interaction between cognitive traits such as cognitive inflexibility (Zmiigrod et al., 2019) or anxiety (Friedman & Thayer, 1998), which are associated with narrower semantic associations (Mikulincer et al., 1990), and a homogeneous political context that trains one’s network of semantic associations.

Third, there might be a polarization of emotional experience since individual differences in life history and psychological need can yield quite different emotional responses to the same political events (Tashjian & Galván, 2018). In one study, partisans watching a 2012 presidential debate between Barack Obama and Mitt Romney had such distinct facial expressions of emotion that voter preference could be predicted from facial expressions alone with over 73% accuracy (McDuff et al., 2013). Given the potential role of emotions in making sense of one’s experiences and deciding on appropriate behavioral responses (Damasio, 1996; Siegel et al., 2018), emotion may play a significant role in construing a polarized subjective experience of political reality (Hatemi et al., 2013; Lieberman et al., 2003; Schreiber & Iacoboni, 2012). Emotions may contribute to political polarization when traits such as cognitive inflexibility cause us to respond in an emotionally inflamed manner (e.g., moral outrage; Goodwin et al., 2009). This can happen when confronted with information that challenges our dearly held preexisting beliefs (e.g., sacred values; Tetlock, 2003), which have been built up by prolonged exposure to a homogeneous political context.

Two central hurdles impede discovering where in the psychological processing pipeline polarization arises and how distinct processing steps might influence one another to create a polarized experience of the political world. First, it is difficult to measure the various psychological processing steps independently, particularly when relying on self-report measures that effectively represent the joint outcome of all steps combined. Second, it has been an enduring challenge to give participants in the laboratory a naturalistic political experience that is strong and rich enough to activate polarization at any relevant processing step. With more traditional cognitive and contextual approaches to polarization, these issues have been less pressing. Cognitive political psychology can index individual differences in cognitive traits using well-vetted experimental tasks (Rollwage et al., 2019), and contextual political psychology can access properties of political environments by looking up voting records or web scraping social media channels (Bakshy et al., 2015; Brady et al., 2017).

There are some behavioral and cognitive science techniques that can address these measurement issues in political psychology. For example, coding face movement and body posture can reveal unfolding emotional states without the need for explicit self-report (Cheong et al., 2020). Skin conductance and facial electromyography also yield information about the strength and valence of emotions in real time (Bakker et al., 2020), and eye tracking provides a window into the attentional focus of the participant (Schmuck et al., 2019). Together, these metrics have made great strides in uncovering the (un)conscious psychological processes that make up political cognition (e.g., Bakker et al., 2020; Petersen et al., 2015; Renshon et al., 2015). Like with any measurement, however, these behavioral and physiological measures also suffer from a few limitations, namely that they often reflect downstream effects of multiple psychological processing steps, such as when skin conductance levels are driven both by attendance and emotional response to a stimulus. Both electroencephalography and functional MRI (fMRI) can overcome these limitations to shed light on the semantic processing of political language—a psychological step that is less accessible to the measurements described above. Put another way, key phenomena such as the subjective perception of the political world remain hidden to the relatively low-dimensional measures in the toolkit of the cognitive–behavioral scientist. Given these limitations, neuroimaging techniques can provide a powerful complement to behavioral and physiological measures of political cognition, aiding in the triangulation of distinct psychological processes that contribute to political polarization.

Although neuroimaging can circumvent the issue of biased self-report by directly tapping into ongoing cognitive processes (Jost et al., 2014), there are limits to traditional brain imaging approaches for few reasons. First, conventional techniques assume that the neural processes of interest are shared across all experimental participants, meaning that one can average neural signals over a group of participants. Yet the very crux of polarization lies in differences in psychological response between individuals. Second, these differences may not manifest in broad changes of neural activity in a given region—which is what traditional neuroimaging methods are sensitive to—but rather in more complex patterns of rising and falling neural activity within a region. In other words, it’s not about documenting where but how the brain is processing political stimuli. Third, traditional neuroimaging approaches use repetitive and highly static experimental tasks, which do not do justice to the complex and ambiguous nature of political events that give rise to political polarization in the first place.

These challenges can be met by capitalizing on several recent developments in fMRI and other brain imaging techniques. One important methodological step forward is that we can now decode subtly different perceptual experiences (e.g., viewing animate vs. inanimate objects; Kriegeskorte et al., 2008) by measuring spatial patterns of neural activity using multivariate representational similarity analysis (RSA; Kriegeskorte et al., 2006; Norman et al., 2006). In contrast to traditional brain mapping techniques, RSA can uncover not just where a stimulus is encoded but also how it is being encoded. For instance, pictures of cats and fish are activate the same brain regions, but activity patterns within these regions can tease apart the representation of the “fish” category from that of the “cat” category. If a researcher now presents a picture of a dog, RSA can demonstrate that the neural response to this animal is more similar to “cat” than to “fish” (Kriegeskorte et al., 2008). This reveals how the brain represents animals in this case, where mammals and fish are treated as two separate categories. One interesting possibility is that the neural encoding of concepts is more rigid in people with low cognitive flexibility—for instance, by organizing concepts into distinct categories with sharp (rather than fuzzy) boundaries (Mikulincer et al., 1990). Translating this into the political domain, certain cognitive traits shaped over years of exposure to partisan news may analogously give rise to a specific and rigid organization of political knowledge. If this is true, RSA may be able to measure this by detecting the similarity between neural activation patterns elicited by pictures of political candidates (Young et al., 2014) and politically relevant events (e.g., immigration)—which would detail a possible mechanism of polarization. Next, by computing how similar neural representations are between participants (an analysis known as intersubject RSA; P. H. A. Chen et al., 2020; Finn et al., 2020; van Baar et al., 2019), it is possible to test whether...
cognitive traits, contextual influences, or both drive the organization of political knowledge.

An important limitation of RSA is that distinct parts of the brain differ in how sensitive and reliable their neuroimaging signals are. As just a few examples, some brain regions are more sensitive to a particular stimulus category such as faces (Haxby et al., 2001; Kanwisher et al., 1997), certain brain regions are known to represent different combinations of features of the same stimuli (Ahlheim & Love, 2018; Badre et al., 2020), and fMRI pattern reliability is lower in the frontal cortex than in other cortices (Bhandari et al., 2018). These regional differences in sensitivity and reliability have two important implications. First, if one region shows distinct representations of political stimuli while another does not, this does not imply that the latter region is not implicated in polarization; it may simply be less sensitive to the presented stimuli. Second, it is unclear whether the subtle processing differences of interest to political neuroscientists can be reliably recovered from brain activity patterns—and if so, in which region. These caveats must be addressed through experimentation. If successful, pattern-based analyses that measure how political content is encoded may illuminate how cognitive and contextual factors interact to mold the subjective perception of polarizing information in our brain.

Researchers have also begun to leverage the temporal trajectory of the brain response (i.e., activity time courses) to reveal participants’ evolving and subjective experience of a stimulus. The most prominent approach is to compute the intersubject correlation (ISC) of signal time courses in a given brain region between individuals (Finn et al., 2020; Hasson et al., 2004; Nastase et al., 2019), which can reveal differences in psychological experience. For instance, one study showed that after being exposed to different prior knowledge, listening to a related story in the brain scanner elicited distinct time courses of neural activation that reflected differences in how the story was being interpreted (Yeshurun et al., 2017). There are several key advantages of an ISC approach. First, a polarizing experience can now be simulated more effectively in the lab. Activity time courses can be time locked to a complex, continuous stimulus such as a video or audio clip (Finn et al., 2018; Leong et al., 2020; van Baar et al., 2020; Yeshurun et al., 2017), which means we no longer have to rely on static, repetitive stimuli. This is crucial for studying polarization, which is known to arise from consuming rich, ambiguous political information in a partisan context (Pomerantsev, 2019). Second, instead of averaging the recorded brain signal across all participants, ISC s leverage the comparison of brain responses between individuals. This makes it possible to test how differences in subjective experience—a hallmark of polarization—arise between individuals (van Baar et al., 2020). Participants come into the lab with preexisting differences in their prior knowledge, epistemic needs, and so forth, allowing us to test which of these individual differences drive differences in the neural processing of political information. Moreover, since the ISC method retains the spatial specificity of fMRI, we can uncover whether the influence of cognitive traits, contextual factors, and their interaction takes hold across distinct steps in the psychological processing pipeline (attention, semantics, etc.). Third, continuous behavioral and physiological measures such as skin conductance, eye movements, and facial expressions can be time locked to the neural response (Chang et al., 2018), allowing researchers to pin down with even more precision which psychological processes (e.g., attention and emotion) are associated with the observed differences in neural response to subjective experience of political video stimuli.

Unlike standard imaging approaches, similarity-based techniques including ISC and intersubject RSA allow the researcher to arbitrate between competing psychological hypotheses about polarization. For example, are biased brain responses to political advertisements best explained by individual differences in ideology (Levendusky, 2009), by differences in epistemic needs such as intolerance to uncertainty (Jost et al., 2003; Zmigrod, 2020), or by both factors in interaction (Hatemi & McDermott, 2016)? Such a question would be difficult to answer using traditional general-linear-model-based analysis of fMRI data as this can only detect increases or decreases between participants in their average BOLD response. In contrast, intersubject RSA and ISC only require researchers to predict which pairs of subjects will be more or less similar to each other in their response pattern, allowing much more flexibility in what the response patterns should look like (Finn et al., 2020; Nastase et al., 2019; van Baar et al., 2019). The observed intersubject similarity or correlation in brain response can then be modeled using any task or survey measure that produces individual differences, ranging from voting history to cognitive batteries to characterizations of participants’ social media networks (Bayer et al., 2018). This versatility in linking behavioral and neural metrics at the intersubject level allows the researcher to integrate levels of analysis that have been traditionally disjointed in the study of political polarization. This includes testing how key contextual and cognitive influences interact to drive polarized perception. For instance, we could take the recorded neural activity of participants in response to a political debate and model the observed ISCs based on three predictors: political ideology, need for closure, and the interaction between these two. Adding these three predictors to a model of ISCs (G. Chen et al., 2017; G. Chen et al., 2019) allows us to test the importance of cognitive-contextual interactions relative to the main effects of cognition and context. We can even use intersubject functional connectivity to reveal whether polarized processing in one region of the brain cascades into polarized processing in the next (Leong et al., 2020; Simony et al., 2016). This suite of
similarity-based analyses can spur new insight into how polarization arises from cognition-context interactions at various psychological processing steps—which jointly construe the polarized experience of reality conjectured by Hastorf and Cantril 70 years ago.

Even with these new developments in cognitive neuroscience techniques, neuroimaging is not a panacea for political psychology—its merits are useful but also limited. For example, there are many well-known statistical and inferential pitfalls that plague neuroimaging, including false positives, circularity in analysis (“double dipping”), and reverse inference (when researchers infer a psychological function from an observed pattern of brain activation; Eklund et al., 2016; Poldrack, 2006; Vul et al., 2009). Perhaps more to the point, imaging methods are most useful when inferring processes that are not attributable to a discrete behavioral analog or when they seek to uncover the neural underpinnings of well-defined psychological processes using hypothesis-driven designs (Niv, 2020). Ideally, neural data can arbitrate between multiple competing hypotheses for the implementation of a psychological process, each captured by a precise cognitive model (e.g., Hampton et al., 2008; Haxby et al., 2011; van Baar et al., 2019). Imaging methods applied to revealing the representations and computations of political polarization can—we hope—achieve this.

In the long run, these multimethod approaches may increase the construct validity of key concepts in political psychology by providing more opportunities for testing convergent and discriminant validity. As illustrated by various examples above, correspondence between behavioral and neural measures can provide a sharper and more mechanistic definition of central concepts like affective polarization. For example, should behavioral reports of affective polarization mirror the physiological data obtained from the brain, it would provide convergent validity. On the other hand, a failure to observe a mapping between behavior and brain data would suggest that the effects be interpreted with caution or that there might be another latent variable at play.

A Path Forward for Using Context Cognition to Understand Polarized Perception

With the advent of recent cutting-edge physiological measurement methods, candidate contributors to political polarization can be rigorously compared in a naturalistic setting. Imagine if we reran the Dartmouth-Princeton football experiment (Hastorf & Cantril, 1954) to test which psychological processing step was most crucial in driving the supporters’ biased judgments of the violence on the field. We could use intersubject neuroimaging approaches to measure the polarization of brain responses in functionally distinct neural networks and yoke those to a variety of psychological and physiological measurements, including visual attention (e.g., eye tracking), semantic understanding (e.g., semantic fluency tasks, Halpern & Rodriguez, 2018; language models, Huth et al., 2016), and emotional responses (e.g., galvanic skin response). When combined with individual difference measures of cognitive traits (e.g., epistemic needs; Zmigrod et al., 2019) and contextual factors (e.g., social media use and prior news consumption about the game), we could drill down into whether Princeton or Dartmouth students were attending to certain plays on the field, feeling differentially aroused during specific violent segments, or interpreting the words of their fellow spectators in dissimilar ways. With such a data set in hand, we could test whether polarization begins at the level of visual attention, physiological arousal, semantic representation, or all of these simultaneously—and characterize how these levels of processing impact one another.

In a world where advertising companies such as Cambridge Analytica are already exploiting voters’ cognitive vulnerabilities for political gain by tailoring information through social media bots, smear campaigns, and fake news (Kosinski et al., 2013; Pomerantsev, 2019; Vosoughi et al., 2018), understanding how the human mind becomes polarized has never been more pressing. A combined cognitive-contextual approach to political psychology can illuminate how individuals become entrenched in a polarized political landscape by unpacking the psychological mechanisms of polarization. This effort may eventually have broad-ranging policy implications, sparking new interventions for more cohesive societies and effective democracies. Leveraging brain imaging to directly test the psychological mechanisms of polarized perception will prove to be invaluable in this endeavor.

References


This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter dur-
Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political
Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political
Baron, J., & Jost, J. T. (2019). False equivalence: Are liberals and conserv-
vatives in the United States equally biased? Perspectives on Psychologi-