

# **Bounded Reflectivism and Epistemic Identity**

**Nick Byrd**

College of Arts and Letters, Stevens Institute of Technology, Hoboken, New Jersey, USA

## **Correspondence:**

Nick Byrd, Stevens Institute of Technology, Peirce 308, 1 Castle Point Terrace, Hoboken, NJ 07030, USA

[byrdnick.com/contact](http://byrdnick.com/contact)

## **Funding**

This research was supported by (a) an appointment to the Intelligence Community Postdoctoral Research Fellowship Program at Carnegie Mellon University and Stevens Institute of Technology, administered by Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Office of the Director of National Intelligence, and (b) both the Department of Philosophy and the Graduate School at Florida State University.

## **Acknowledgments**

My thanks to James Beebe, Mike Bishop, Otto Bohlmann, Rhys Borchert, Paul Conway, Simon Cullen, Caleb Dewey, Kevin Dorst, Anders Ericsson, Reza Hadisi, Mike Huemer, Josh Knobe, Matt Jones, Jack Justus, Anabella Lichter, Andrew Lichter, Edouard Machery, Heather Maranges, Gordon Pennycook, Caleb Reynolds, John Schwenkler, Michael Slote, Houston Smit, Brian Talbot, Marshall Thompson, Jason Turner, Jay Van Bavel, Cassidy Velasquez, Jonathan Weinberg, Chase Wrenn, and anonymous referees. Also to audiences at the Australasian Society for Philosophy and Psychology conference (2018), Southern Epistemology Conference (2018), and University of Arizona Department of Philosophy Colloquium (2021).

*In press at [Metaphilosophy](#) (2022). Accepted November 2021. See [byrdnick.com/cv](http://byrdnick.com/cv) for more.*

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

### **Abstract**

Reflectivists consider reflective reasoning crucial for good judgment and action. Anti-reflectivists deny that reflection delivers what reflectivists seek. Alas, the evidence is mixed. So, does reflection confer normative value or not? This paper argues for a middle way: reflection can confer normative value, but its ability to do this is bound by such factors as what we might call *epistemic identity*: an identity that involves particular beliefs—for example, religious and political identities. We may reflectively defend our identities' beliefs rather than reflect open-mindedly to adopt whatever beliefs cohere with the best arguments and evidence. This *bounded reflectivism* is explicated with an algorithmic model of reflection synthesized from philosophy and science that yields testable predictions, psychometric implications, and realistic metaphilosophical suggestions—for example, overcoming motivated reflection may require embracing epistemic identity rather than veiling it (à la Rawls 1971). So bounded reflectivism should be preferred to views offering anything less.

### **KEYWORDS**

cognitive science, epistemology, value theory, philosophy of mind, dual process theory, deliberative democracy, reflective equilibrium, reflective endorsement, reflection, cognitive reflection test, bounded rationality, political polarization

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

[A]ll men of reflection . . . are either free from erroneous prepossessions, or can divest themselves of them. (Hamilton, Madison, and Jay 1787)

Even advanced reflection and training does not insulate one from illusion: . . . physics graduate students and postdoctoral researchers still experience the characteristic cognitive-perceptual illusions of naïve “impetus theory” physics. (Nagel 2012)

We endorse or reject our impulses by determining whether they are consistent with the ways in which we identify ourselves. . . . You are a mother of some particular children, a citizen of a particular country, an adherent of a particular religion. . . . And you act accordingly—caring for your children because they are your children, fighting for your country because you are its citizen, refusing to fight because you are a Quaker, and so on. (Korsgaard 1996)

You agreed to cover the tip for lunch with your friend. So as you stand up from the table, you put down an amount of money that feels right. Your friend glances at the money and appears surprised. So you do some calculation in your head, realize that you forgot to factor in your friends’ portion of the bill, and add some money to the gratuity. On the way out of the restaurant, your friend asks you about your political party’s latest scandal. You immediately play defense, rehearsing various rationalizations of the scandal. After your monologue, your friend recalls that you criticized the opposing party for the same kind of scandal in the last election.

This story reveals a puzzle about reflective reasoning: reflection often helps, but reflection can also hinder our reasoning. Reflection is crucial for double-checking and correcting our reasoning, but reflection can also be used to rationalize bad reasoning when we feel that our cherished beliefs are under attack. This puzzle about reflection is also manifested in scholarly

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

debates about reflection. Reflectivists claim that reflection is crucial for good judgment and action (e.g., Epstein 1994; Klein 1998; Korsgaard 1996; Sosa 1991), but anti-reflectivists admit that reflection can lead us astray (e.g., Doris 2015; Kornblith 2012). So what gives? Does reflection help or does it hinder reasoning?

To address this puzzle about reflection, I offer a middle way between reflectivism and anti-reflectivism. Reflection does not necessarily make our reasoning good or bad. Rather, reflection is a tool. Its utility depends largely upon other factors, such as our goals. For instance, when someone wants to defend the beliefs with which they most identify, reflection may not satisfy nonpartisan epistemic standards. In other words, reflection's quality can be bound by the beliefs that we tie to our identity. The remainder of the paper explains, argues for, and elaborates on how this notion of epistemic identity leads to the view that I call *bounded reflectivism*. Section 1 lays out the theory and evidence about the normative value of reflection. Section 2 introduces the notion of epistemic identity and its potential impact on reflection. Section 3 synthesizes the bounded reflectivist model of reflection from the prior sections and relevant literature. Section 4 outlines the philosophical and scientific implications of bounded reflectivism. And section 5 offers some concluding remarks.

### 1 | WHAT'S SO GREAT ABOUT REFLECTION?

Reflective reasoning is a mainstay in the history of ideas. John Doris puts it this way: “A preoccupation with reflection is, arguably, the Western philosophical tradition’s most distinctive feature, in both historical and contemporary contexts” (2015, chap. 2). This concept of reflection is familiar. A famous first-person description of reflection comes from Christine Korsgaard: “I find myself with a powerful impulse to believe. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn’t dominate me and now I have a problem. Shall I believe?” (1996, sec. 3.2.1). Reflection is a thriving area of research in cognitive science as well. Well-known titles like *Thinking Fast, and Slow* (Kahneman 2011) have popularized much of this research on reflective and unreflective reasoning (De Neys 2018; Pennycook 2018). So reviewing the literature on the value of reflective reasoning involves both philosophical theory and scientific investigation.

### 1.1 | Theory: Reflectivism and anti-reflectivism

Reflective reasoning is said to be more deliberate (that is, less automatic) and more consciously represented than unreflective reasoning (e.g., Byrd 2019; forthcoming; Shea and Frith 2016). So tests of reflection have been designed to lure us into automatically accepting a believable conclusion that—if we stop and consciously represent all the relevant information—we can realize is incorrect. For instance, “If we know only that flowers have petals and that roses have petals, can we conclude that roses are flowers?” (adapted from Byrd and Conway 2019). Our prior belief that roses are flowers lures us into automatically accepting the inference. If, however, we deliberately stop ourselves from accepting this inference and think about the logical structure of the question (that is, if we reflect), then we can realize that the inference is fallacious: it affirms the consequent. In addition to operationalizing “reflection,” reflection tests show why reflection can be valuable. Reflection can be instrumental in overcoming misleading heuristics and biases, such as the tendency to accept conclusions that align with our prior beliefs (Byrd et al. forthcoming). Of course, the value of reflection is a topic of debate.

Many philosophers take reflection to be crucial for obtaining various intellectual goods. We can call such philosophers reflectivists (Ferrin 2017). Consider a selection of reflectivism’s history.

- “Reflective persons” were said to “eliminate error from our [conflicting] moral intuitions [via] appeal to general rules or formulae” (Sidgwick 1874).
- “Reflective equilibrium” was said to be necessary to discern and/or justify principles of logic and justice (Goodman 1983; Rawls 1971).
- “Reflective agency” was said to be important to understanding human action (e.g., Kennett and Fine 2009; Velleman 1989; Wallace 2006).
- “Reflective knowledge” was said to be a distinctive capacity of humans that is necessary to understand our beliefs in context and “how they come about” (Sosa 1991).
- “Reflective endorsement” was said to be necessary for morality to have normative force (Korsgaard 1996).

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

- “Reflective scrutiny” was said to be necessary for evaluating our ethical outlook from within—as opposed to evaluating neutrally, from the outside, or by merely re-expressing itself (Hursthouse 1999).

There are probably more instances of reflectivism in philosophy (see Doris 2015). While an exhaustive catalog of each instance is a valuable historical project, the goal of this section is just to introduce reflectivism’s ongoing and wide-reading presence in philosophy.

Opposing reflectivism, of course, is anti-reflectivism. Anti-reflectivists argue that reflection cannot do or be what reflectivists think. Consider an example of anti-reflectivism: “Reflection, by and large, does not provide for greater reliability. It does not, by and large, serve to guard against errors to which we would otherwise be susceptible. It does not, by and large, aid in the much-needed project of cognitive self-improvement. It creates the illusion that it does all of these things, but it does not do any of them” (Kornblith 2012, 26).

Additional examples of anti-reflectivism include arguments that reflection cannot be a virtue (Dreyfus 1986), that “reflection on our beliefs and decisions distorts our view of our own mental processes” (Kornblith 2019, 19), and that reflection does not give us the self-knowledge that many reflectivists imagine (Doris 2015).

So who is right? The reflectivists or the anti-reflectivists? In my view, both reflectivists and anti-reflectivists get *something* right. Reflectivists are right to think that reflective reasoning delivers some of the normative value that reflectivists care about. Anti-reflectivists are right, however, to think that reflectivists seem to be unduly optimistic about reflection: in some circumstances reflection lacks the proposed normative value and—worse—in other circumstances reflection can produce normative disvalue. By considering the evidence for these claims, we can begin to see the shape of the middle way that this paper develops.

### 1.2 | Evidence: Good and bad reflection

Public discourse in the United States often focuses the ways in which online fake news can be weaponized to undermine democracy. Reddit’s CEO announced cooperation with federal investigations of the dissemination of fake news on its website and concluded with the belief that “the biggest risk we face as Americans is our own ability to discern reality from nonsense”

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

(Huffman 2018). Part of the worry was that reasoning is politically partisan. Some researchers find evidence for such partisan reasoning. For example, multiple studies of more than a thousand people found that political conservatives were far more likely to consider an action morally wrong if it was performed by a left-wing than by a right-wing person:  $b = 0.33-0.41$ ,  $SE = 0.04-0.58$ ,  $t = 6.87-9.31$ ,  $p < 0.001$  (Everett et al. 2021). And in a meta-analysis of more than fifty studies involving more than eighteen thousand people, liberals and conservatives rated politically congenial information as more valid, of higher quality, or more acceptable than politically uncongenial information:  $k = 51$ , meta-analytic  $r = 0.245$ , 95% CI [0.208, 0.280],  $p_r < 0.001$ ,  $Q_w = 307.96$ ,  $p_w < 0.001$ , Tau = 0.120 (Ditto et al. 2018). Yet even though reflection correlated positively with politically biased news evaluation among more than four hundred U.S. adults,  $0.12 < r < 0.2$ ,  $0.036 < p < 0.001$ , the same experiment found that reflection also correlated positively with better news discernment,  $r = 0.26$ ,  $p < 0.001$  (Calvillo and Smelter 2020, Experiment 2). It is mixed evidence like this that makes one wonder whether reflection helps or whether it hinders our reasoning.

*Reflection as a solution.* In two studies of more than eight hundred participants, people who scored better on certain reflection tests (Frederick 2005; Thomson and Oppenheimer 2016; section 3.1 below) were significantly more likely to correctly estimate the accuracy of fake news—Study 1:  $r = -0.3$ ,  $p < 0.001$ ; Study 2:  $r = -0.26$ ,  $p < 0.001$ —and significantly less likely to share fake news—Study 2:  $r = -0.19$ ,  $p < 0.01$ —even when the source of the news was removed and when headlines aligned with their partisan identity (Pennycook and Rand 2019a). Across two other studies of about two thousand participants, such reflection predicted more reliance on mainstream news sources over hyperpartisan and fake news—Study 1:  $\beta = 0.16$ ,  $p < 0.0001$ ; Study 2:  $\beta = 0.15$ ,  $p < 0.0001$  (Pennycook and Rand 2019b). In other words, reflective reasoning was associated with more desirable reasoning in everyday and seemingly high-stakes contexts such as political reasoning. These findings suggest that reflective reasoning can be part of the solution to problems with our ability to discern reality and with non-sense.

*Reflection as ineffective.* Unfortunately, reflective reasoning is not a panacea. Consider the illusory truth effect. Multiple studies of more than a thousand people found that encountering false information repeatedly made people more likely to believe it regardless of their performance on reflection tests—Studies 1, 2, and 5 meta-analytic  $r = -0.01$ , 95% CI [-0.09, 0.07]; Studies 5, 6, 7

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

meta-analytic  $r = -0.05$ , 95% CI  $[-0.10, 0.01]$  (De keersmaecker et al. 2019). These findings suggest that some reasoning problems might be immune to reflection.

*Reflection as a problem.* In fact, reflection might make matters worse. Multiple experiments found that people who are more likely to reason reflectively are also more likely to reflect in ways that serve their partisan identities. For instance, Democrats in the United States are far more likely to endorse the claim that there is “solid evidence” of global warming than Republicans in the United States (Pew Research Center 2017). Also, when participants were told that open-minded people who accept climate change are more likely to get the correct answers on the Cognitive Reflection Test (CRT), right-leaning participants were less likely to report that the CRT was valid, while left-leaning participants were more likely to report that the CRT was valid (Kahan 2013). Crucially, this partisan evaluation of the CRT increased (rather than decreased) among more reflective participants,  $r = -0.3$ ,  $p < 0.01$  (Kahan 2013).

In another experiment, participants interpreted fictional studies. More reflective participants were more likely to correctly interpret the findings of studies about nonpoliticized topics. Alas, when interpreting studies about politically salient topics like gun control policy, both left-leaning and right-leaning individuals were more likely to misinterpret evidence that challenged their policy preferences (Kahan et al. 2017; Connor et al. 2020; cf., Persson et al. 2021). Once again, this partisan evaluation of evidence was more dramatic among more reflective participants,  $r = 0.54$ ,  $p < 0.05$  (Kahan, Peters, et al. 2017, Figures 3 and 4). In more recent studies, reflecting on initial judgments about real and fake news improved peoples’ truth discernment, but people still (upon reflection) judged politically concordant news to be more accurate than discordant news,  $b = -0.21$ , 95% CI  $[-0.34, -0.07]$ ,  $p = 0.003$  (Bago, Rand, and Pennycook 2020). These findings suggest that while reflection can help reasoning in some cases, reflection can also hinder our reasoning.

Overall, these data suggest that reflection can be part of the problem, insufficiently effective, or else part of the solution. These mixed results are similar to what launched the bounded rationality literature, which attempts to explain how rationality can retain its normative value while being less perfect than originally suggested (Wheeler 2018). So the current project involves explaining how some of the normative value that reflectivists attribute to reflection can be retained even if—as anti-reflectivists have argued—that value is bounded in ways that reflectivists may not



have originally suggested (see also Stephens and Tjøstheim 2020). One specific bound on reflectivism is a kind of partisanship that I call *epistemic identity*.

### **2 | EPISTEMIC IDENTITY**

Korsgaard (1996) famously discusses the roles of practical identity in reflection. Another form of identity is what I call epistemic identity: the phenomenon of treating certain beliefs as part of one's identity. Suppose that I identify with a religion. If you criticize some aspect of my religion, then I might reflectively defend my religious beliefs rather than dispassionately submit to the best arguments and evidence. In short, I might prioritize my epistemic identity over other epistemic goods. Or suppose that you identify with a particular political party—one that explicitly codifies its ideological commitments in a party platform that is recited in its public speeches, advertisements, and so on. In other words, you identify not only with the party but also with its values and beliefs. In this case your political identity is an epistemic identity.

We have already encountered evidence of an effect of epistemic identity on reflection. There is also evidence, however, of the impact of epistemic identity on reasoning more generally. For instance, judgments about evidence of global warming are more correlated with self-reported political ideology (“left” versus “right”) in the United States than in any of the other twenty-five Western, developed countries tested,  $d > 0.4$  (Hornsey, Harris, and Fielding 2018, Figure 1). But the United States is not alone. Italian citizens have also been found to consider politically concordant news to be more plausible (Vegetti and Mancosu 2020). Given these and similar data, it is unsurprising that identity has become an increasingly common part of reasoning research.<sup>1</sup>

#### **2.1 | Proximal and distal effects of epistemic identity**

We can distinguish between proximal and distal influences on action (Mele 2008, chap. 2). Likewise, epistemic identities can have proximal and distal influences on our reflective reasoning. Epistemic identities have proximal influences on our reflective reasoning when they determine how our reflective reasoning proceeds in any given moment. For example, you might be familiar

---

<sup>1</sup> E.g., Brandt and Slegers 2021; De Cruz 2020; Funkhouser 2020; Oyserman and Dawson 2019; Roets et al. 2020; Strohminger 2018; Van Bavel and Pereira 2018.

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

with an outstandingly opinionated and uncivil person who is quick to share his pet beliefs and immediately dismisses any opposing evidence and arguments—sometimes referred to anecdotally as an “obnoxious uncle” (Lynch 2018). Of course, partisanship can influence reflection in more distal ways. For example, some have argued that philosophers may be implicitly influenced by epistemic identity, thereby shaping the broader philosophical discourse over time (Peters 2019).

*Distal effects of epistemic identity.* Epistemic identities can have subtle, longitudinal influences on not only reasoning in general but also reflective reasoning in particular. Our epistemic identities can influence what we seek, what we attend to, what we perceive, and thereby what we remember, whom we listen to (Levy 2019; Rini 2017), and how we reason (Kahan 2016). So when we represent our memories consciously in order to reason about them deliberately, we are working with systematically biased priors. For example, we may be more likely to seek out, attend to, perceive, and remember the successes and unfair criticisms of our own group than competing groups *and* less likely to seek out, attend to, perceive, and remember the failures and level-headed criticisms of our group than competing groups (Derreumaux, Bergh, and Hughes 2020; Carlson et al. 2020). This is why one epistemic difference between two groups can result from multiple polarizing processes from each group (Talisso 2019). This is also part of the reason why providing people with more information does not always reduce polarization (Gershman 2018; Lee et al. 2020): polarization can occur not only while *receiving* information (à la Lord, Ross, and Lepper 1979) but also while *remembering* it.

*Proximate effects of epistemic identity.* The distal influences of epistemic identity can result in proximate impacts of epistemic identity (Bakker, Lelkes, and Malka 2019; Van Bavel and Pereira 2018). This identity-based reasoning can also explain why many people follow their leaders more than they follow logic: that is, they adopt their leader’s policies even when doing so deviates from the logical conclusions of their own professed ideals, principles, and values (e.g., Kinder and Kalmoe 2017; Rathje, Shariff, and Schnall 2020). Further research might reveal if identity-driven thinking can also explain the opening puzzle: how reflection’s impact may change from helpful or hurtful from one moment to the next. One way to explain this puzzle is to admit that epistemic identity is more likely to negatively influence reflective reasoning in adversarial contexts than in open-minded contexts (Groenendyk and Krupnikov 2020).

## 2.2 | If you can't beat epistemic identity, embrace it

Some have suggested that the solution to the problem of epistemic identity involves imagining someone else's identity (Hannon 2020). Others have suggested that we should activate a superordinate identity (Van Bavel and Pereira 2018). My suggestion is a combination of these suggestions: to overcome the undesirable impacts of epistemic identities, we will need to appeal to *shared, superordinate* epistemic identities.

Consider an example. When a colleague and I are reflectively doubling down on a political disagreement, then we might do well to stop to ask, “Although we disagree, what *should* we think about this *as scientists*?” Because we both identify as scientists and we agree about the standards of scientific reasoning, undesired polarization may be less likely when we reflect *as scientists* who share epistemic standards than when we reflect as political partisans employing politically congruent epistemic standards.<sup>2</sup>

Depending on the context, shared identities need not be shared by everyone. They may only need to be shared by those involved in the undesirably polarized disagreement. For instance, within a politically divided country, appealing to shared national identity—as opposed to, say, a globalist or international identity—might lead to productive agreement (e.g., Talaifar and Swann 2018; Bavel et al. 2020). As the scope of discourse expands to include stakeholders in other countries or groups, however, appealing to broader identities may be required to overcome polarization and divisiveness.

So epistemic identity may not be only a problem; it may also be part of the solution. In fact, some data confirm this. For instance, when U.S. Republicans' party identity was threatened by questions about the deterioration of the Republican party and then they were told that a majority of Republicans agree about global warming and are taking action to combat climate change, they were significantly more likely to endorse and care about climate change than if they were merely asked to be even-handed in their evaluation of climate change:  $t(1951) = 2.29, p = 0.02, d = 0.10$  (Bayes et al. 2020). This suggests that threatening one's epistemic identity can lead not only to undesirable judgments but also to desirable outcomes, depending on the reasoning context. So when we cannot beat epistemic identity, we might need to embrace it.

---

<sup>2</sup> E.g., Kahan, Landrum, et al. 2017; Sikimić et al. 2020; Weisberg et al. 2020; West and Iyengar 2020.

### **3 | THE BOUNDED REFLECTIVIST MODEL OF REFLECTION**

We now have the psychological construct, the evidence, and one of the mechanisms motivating the bounded reflectivist model of reflection. Reflection is said to be deliberate and conscious (Shea and Frith 2016); reflection can deliver goods that reflectivists seek, but reflection can also hinder our reasoning in ways that anti-reflectivists have suggested sometimes as a result of—among other things—epistemic identity. The bounded reflectivist model of reflection that I have in mind, however, is based on a more complete account of reflection, including the triggers of reflection, the outcomes of reflection, and the possible psychological paths between them.

#### **3.1 | Visualizing the bounded reflectivist model**

To make it easier to follow the written explanation of the model, the bounded reflectivist model is visualized algorithmically in Figure 1. As I see it, there may be multiple (albeit finite) ways to arrange these processing steps and decision nodes in this model of reflection that remain compatible with the account of reflection developed in this paper. For instance, the order of the triggers of reflection (outside the large box labeled “reflection”) and the steps of reflection (inside the large box labeled “reflection”) could be slightly rearranged without violating bounded reflectivism. Moreover, there could be cases in which reasoning involves multiple looping paths through this model—for example, from an intuition to reflection about that intuition, and then to another intuition, and so forth (Nagel 2014).

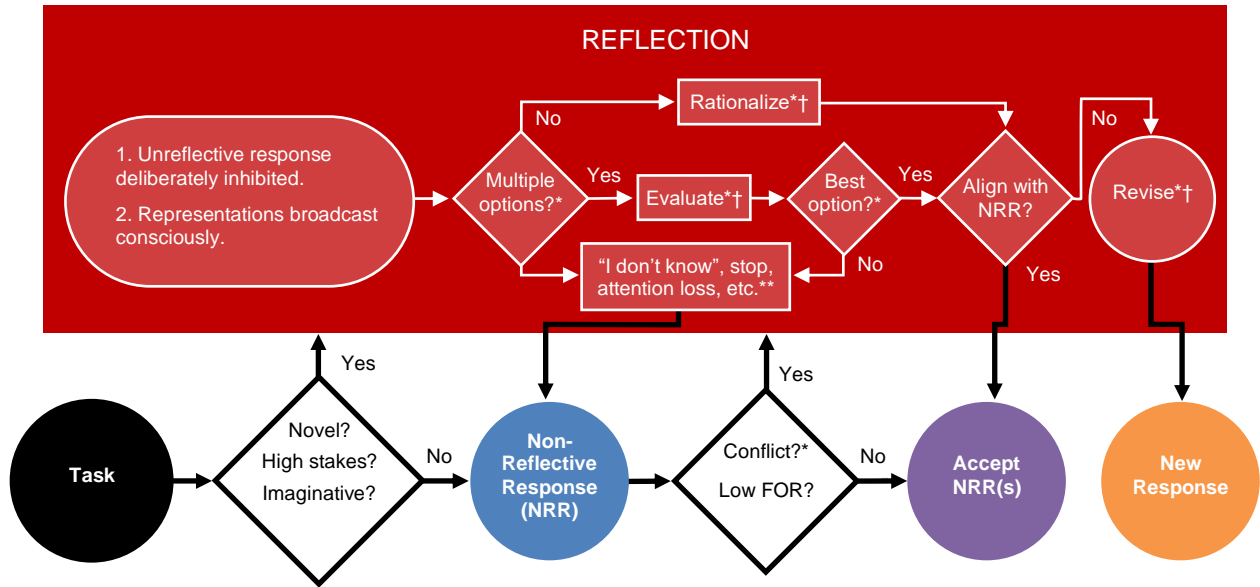


Figure 1. The causes, processes, and outcomes of partisan reasoning, reflective and unreflective. \* Can be influenced by factors like epistemic identity. \*\*Can involve new reasons and/or slightly new response even if same spirit of IR. † If interrupted, null response.

### 3.2 | When reflection is not triggered

The visualization illustrates how tasks can prompt either an autonomous response or reflection. If the task is not novel, high stakes, or imaginative, then reflection will probably not be triggered. Rather, an autonomous response will occur. Whether we accept the autonomous response or reflect on it will depend on whether we detect conflict between the autonomous response and some other seemingly relevant response or whether we have a feeling of rightness about the autonomous response. If no conflict is detected and one’s feeling of rightness is high, then the autonomous response will probably be accepted.

Of course, our acceptance of an autonomous response might be influenced by epistemic identity. After all, we are more likely to unreflectively endorse conclusions that comply with our prior beliefs, even when doing so is logically fallacious (Janis and Frick 1943).

### 3.3 | Triggering reflection

Some have proposed that reflection is triggered either by a task (Nagel 2012) or by how confident one feels about one's autonomous response.<sup>3</sup> Together, these potential triggers can explain performance on reflection tests. Reflection tests are designed to lure us into unreflectively accepting a particular answer that is, upon reflection, incorrect (Byrd forthcoming). For instance, the original bat-and-ball problem lures people to the \$0.10 answer (Frederick 2005):

A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?

A recent variant of the famous bat-and-ball problem, however, may not (Baron et al. 2015):

A bat and a ball cost 96 cents in total. The bat costs 2 cents more than the ball. How much does the ball cost?

This variant, unlike the original version, does not lure participants toward a particular incorrect answer. So there is no obvious reason to think that the correct response on this question involves deliberately inhibiting, feeling right about, or reflecting on a particular autonomous response.

Some suggest, however, that reflection can also be triggered by task novelty, stakes, or imagination (Nagel 2012). Novelty can explain how reflection could be triggered by the non-lured bat-and-ball problem. While the non-lured problem does not elicit a particular autonomous response like its predecessor, reflection might nonetheless be triggered for reasoners who find this mathematical task novel, high stakes, or imaginative. Of course, according to the deliberate-and-conscious account of reflection, reflection would be occurring only if participants consciously represent some of their reasoning and deliberately prevent themselves from *immediately* accepting answers that spring to mind while consciously reasoning.

---

<sup>3</sup> Koriat 2019; Mercier and Sperber 2017; Pennycook, Fugelsang, and Koehler 2015.

### 3.4 | When reflection is triggered

Once triggered, there are a few steps within the reflection process that could result in multiple different outcomes. Factors such as epistemic identity can influence which outcome occurs. And, crucially, the outcome of reflection can depend on the outcomes that occur at each step.

*One versus multiple options.* The first step of reflection involves considering options. “P seems true, but is it?” “Might Q be true instead?” “What about not-P?” The search for options might end with an empty set of options—for example, “I don’t know”—or else get interrupted, making a reasoner more likely to opt for their unreflective response—for example, in cases of cognitive load (e.g., De Neys 2006, Figure 2; Vonasch 2016, Figure 1). Alternatively, the search can reveal one or more options. If only one option is considered—either because someone reflects in a close-minded manner or for some other reason—then we can reflectively rationalize it (Cushman 2019). If more options are considered—either because someone reflects in a more open-minded manner (Stanovich and West 1997) or for some other reason—then we can reflectively evaluate the options and decide whether there is a best option. Again, epistemic identity can influence the process of reflection—for example, which options are considered and how options are evaluated (Kahan and Corbin 2016).

*Evaluation.* During reflective evaluation, we might find that an autonomous response was, in fact, the best response.<sup>4</sup> In this case, reflective reasoning would result in accepting the autonomous response—even if for new and/or better reasons than we considered prior to reflection—à la “reflective endorsement” (Korsgaard 1996). This reflective evaluation, however, might instead reveal that an alternative response is superior. In this case, reflection would revise the initial unreflective response to a new response. Of course, sections 1 and 3 above remind us that we sometimes reflectively evaluate arguments and evidence under the influence of epistemic identities, especially when we feel that our epistemic identities or the beliefs they entail are under attack.

---

<sup>4</sup> Byrd et al. forthcoming; Bago and De Neys 2017; Raelison et al. 2021.

### 4 | IMPLICATIONS

This algorithmic bounded reflectivist model has implications for both philosophy and science. First, the model operationalizes reflection in ways that can guide scientific inquiry—for example, by proposing empirically trackable ways in which reflection can be triggered and influenced by identity-based thinking. The model also has implications for both scientists' reflection tests and philosophers' views about the normative value of reflection.

#### 4.1 | Measures of reflection

Many researchers have assumed that correct responses on reflection tests involve overcoming a default (that is, autonomous) response—hence, the default-interventionist account of analytic reasoning (Evans 2007; see also Johnson-Laird and Ragni 2019). But when researchers record participants thinking aloud as they complete these measures of reflection, they often find that participants can immediately respond correctly with no verbal or other evidence of a lured response (Szasz et al. 2017; Byrd et al. forthcoming). This suggests that the default-interventionist model of reflection cannot account for all reflection test performance.

The bounded reflectivist model of reflection can, however, explain these immediate and correct responses on reflection tests. In particular, the model illustrates a route to a successful response that does not involve reflection: when a task is low stakes—as when participants are not rewarded or punished for performance on simple arithmetic questions—and when the task is so familiar that it requires no imagination, participants' autonomous responses might be correct prior to reflection.

Also, while real-time—versus post hoc (Byrd forthcoming)—verbal reports of reasoning have revealed that some correct reflection test responses do not involve reflection, which aligns with the bounded reflectivist model, we might wonder if further think-aloud protocol analysis could provide novel tests of the model. For example, think-aloud recordings and transcripts have the potential to falsify bounded reflectivism's proposed triggers of reflection, steps involved in reflection, and how other factors can influence the outcomes of each of these steps (such as identity cues or reduced cognitive load.)



### 4.2 | The normativity of reflection

Recall the reflectivists who argue that reflection is crucial for good reasoning (e.g., Korsgaard 1996; Sosa 1991). These reflectivists think that reflection is always normatively valuable. And that is part of what anti-reflectivists deny (e.g., Kornblith 2012; Doris 2015). Bounded reflectivism accepts many of the anti-reflectivists' negative conclusions about reflection without going so far as to say that the normative value of reflection is an illusion. Hence, bounded reflectivism fulfills the need for a more "sensible reflectivism" between reflectivism and anti-reflectivism (Schwenkler 2018). Bounded reflectivism also provides other normative implications about how to handle epistemic identity and the bounded value of reflection.

*Reflective equilibrium.* Bounded reflectivism's middle way between reflectivism and anti-reflectivism involves admitting that reflection's normative value is contingent. Reflection can confer normative value in some conditions and not in others, depending on factors like epistemic identity. Some have realized that reflection is not enough for reflective equilibrium. For example, some suggest that reflection must occur behind a veil of ignorance to insulate it from the undesirable impacts of epistemic identity (Rawls 1971). Some, however, have questioned the psychological plausibility of ignoring our own identity when reflecting (Levin 1978). The bounded reflectivist model reinvigorates the plausibility of reflective equilibrium with a more realistic suggestion than Rawls's original position: rather than veil epistemic identity, embrace it. Bounded reflectivism suggests that reflective equilibrium is more likely when reflective reasoners are motivated by shared, superordinate epistemic identities than when reflective reasoners are under the influence of narrower, less universal epistemic identities.

*Reflective scrutiny.* Bounded reflectivism's treatment of identity can also address the objectivity problem. Consider how the objectivity problem is manifested in metaphilosophy: when we try to justify our framework from within the framework, we end up merely re-expressing the framework rather than justifying it—for example, when we selectively focus on the desiderata that our framework was specifically designed to satisfy. Some philosophers suggest that reflection is a solution to this problem (Hursthouse 1999). Bounded reflectivism takes a different approach, however, holding that metaphilosophical scrutiny may require a shared, superordinate epistemic identity.

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

Naturally, shared, superordinate epistemic identities cannot achieve *external* objectivity. Such objectivity may be unachievable, however (Carnap 1950), suggesting that the external objectivity problem is not endemic to bounded reflectivism. What shared superordinate epistemic identities can offer is a shared *internal* framework that often *is* achievable—for example, when scientists put aside their other epistemic identities to consider what they ought to think about a puzzle *as scientists*. Without such shared epistemic identity (or identities) to provide a common internal framework, it is not clear how reflection about your view could have optimum normative force with me (and vice versa). So if reflective scrutiny is to work *between* people (and not just *within* a single person), shared epistemic frameworks (such as those provided by bounded reflectivism’s epistemic identities) will be necessary.

*Strategic reliabilism.* Of course, this paper does not consider all the conditions in which reflection will help or hinder our normative goals. Research can enumerate the ways reflection is used toward normative goals and test its efficacy across contexts.<sup>5</sup> The resulting data could feature in a sort of strategic reliabilist account of reflection’s normative value—*strategic reflectivism*: reflection should be deployed in contexts where its benefits have been shown to reliably outweigh its costs (à la Bishop and Trout 2004 and 2008; Stich 1990).

Strategic reflectivism’s admission that reflection can improve our reasoning is more optimistic than the anti-reflectivism according to which reflection “gives us the illusion that we have subjected our beliefs to a rigorous screening that will improve our epistemic position,” when, “[i]n fact, . . . it achieves no such thing” (Kornblith 2019, 4). Similarly, strategic reflectivism’s admission that reflection can produce normatively undesirable outcomes is more pessimistic than the reflectivism that sometimes characterizes normativity as “the ability to survive reflection” (Korsgaard 1996, lecture 9, sec. 5). Hence, a strategic reflectivism based on bounded reflectivism could also provide a middle way between competing views about the normative value of reflection.

## 5 | CONCLUSION

Bounded reflectivism draws on philosophy and psychology to propose a unifying, empirically adequate, testable, and actionable model of reflective reasoning. Nonetheless, not all questions

---

<sup>5</sup> Davies, Ives, and Dunn 2015; Savulescu, Kahane, and Gyngell 2019.

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

about reflection and identity are answered in this paper. For instance, are identity-driven and belief-driven reflection epistemically suspect “motivated reasoning” (Kahan 2016) or merely epistemically rational Bayesian reasoning (Tappin, Pennycook, and Rand 2021; Dorst forthcoming)? How might polarizing identity-driven reflection *improve* inquiry (Gampa et al. 2019; Kitcher 1990)? Further attention to questions like this will certainly advance our understanding and appreciation of reflection and epistemic identity.

There may also be questions about whether reflection could be normatively valuable *independently* of its consequences. Even some of my deontologically minded colleagues admit, however, that they would be surprised if the value of reflection had *nothing* to do with its moral and epistemic consequences. For instance, if reflective reasoning made physician-scientists reliably decrease the quality and longevity of their patients’ lives and the accuracy of their medical theories compared to nonreflective reasoning (for example, passively accepting the decisions of an algorithm), then in these conditions reflection seems *normatively* worse than its alternatives even if the reflection produced *psychological* goods (for example, a sense of autonomy, self-awareness, individual reflective equilibrium, and so on). Likewise, if reflection made people reliably decide to opt out of low-risk, low-cost, convenient medical treatment that is known to drastically decrease not only their own suffering and death but also global suffering and death, but these people ended up consenting to the treatment via less reflective thinking (e.g., a subliminal bandwagon effect), then the less reflective decision seems better than its reflective alternative. Put another way, dismissing reflection seems better than heeding it in these cases—at least partly because of consequences. And this need not be anathema to non-consequentialists. As I say elsewhere, even non-consequentialists have “considered both ‘bad results’ or ‘disastrous consequences’” in their normative theory (Byrd and Conway 2019). So even if there is more to reflection than its consequences, we may not fully disconnect consequences from the normative value of reflection. Further investigation may reveal the merits and demerits of consequence-free views of reflection (à la Littlejohn 2018).

In the meantime, the bounded reflectivist model offers value to both philosophers and scientists. It delivers an algorithmic explication of reflection that unifies philosophical and scientific theories of reflection, makes testable predictions about reflection, guides scientific measurement of reflection, and provides normative suggestions about how to use reflection. One side effect of bounded reflectivism’s benefits is a middle way between the well-traveled paths of

## BOUNDED REFLECTIVISM & EPISTEMIC IDENTITY

reflectivism and anti-reflectivism: reflection can be normatively good—for example, when shared epistemic identities allow for reflective equilibrium—or normatively bad—for example, when reasoners refuse to adjust their commitments to the conclusions with which they identify in the face of compelling arguments and evidence—or even normatively moot in certain (for example, familiar) circumstances in which nonreflective responses are at least as good as reflective responses (à la Rubin et al. 2019). Insofar as other accounts of reflection cannot deliver all these goods, we should prefer the bounded reflectivist model of reflection.

## REFERENCES

- Bago, Bence, and Wim De Neys. 2017. “Fast Logic?: Examining the Time Course Assumption of Dual Process Theory.” *Cognition* 158 (January): 90–109. <https://doi.org/10.1016/j.cognition.2016.10.014>.
- Bago, Bence, David Rand, and Gordon Pennycook. 2020. “Fake News, Fast and Slow: Deliberation Reduces Belief in False (but Not True) News Headlines.” *Journal of Experimental Psychology: General*. <http://dx.doi.org/10.1037/xge0000729>.
- Bakker, Bert, Yphtach Lelkes, and Ariel Malka. 2019. “Understanding Partisan Cue Receptivity: Tests of Predictions from the Bounded Rationality and Expressive Utility Perspectives.” *The Journal of Politics*, December. <https://doi.org/10.1086/707616>.
- Baron, Jonathan, Sydney Scott, Katrina Fincher, and S. Emlen Metz. 2015. “Why Does the Cognitive Reflection Test (Sometimes) Predict Utilitarian Moral Judgment (and Other Things)?” *Journal of Applied Research in Memory and Cognition*, Modeling and Aiding Intuition in Organizational Decision Making, 4 (3): 265–84. <https://doi.org/10.1016/j.jarmac.2014.09.003>.
- Bavel, Jay J. Van, Aleksandra Cichocka, Valerio Capraro, Hallgeir Sjøstad, John B. Nezlek, Mark Alfano, Flavio Azevedo, et al. 2020. “National Identity Predicts Public Health Support during a Global Pandemic: Results from 67 Nations.” PsyArXiv. <https://doi.org/10.31234/osf.io/ydt95>.
- Bayes, Robin, James N. Druckman, Avery Goods, and Daniel C. Molden. 2020. “When and How Different Motives Can Drive Motivated Political Reasoning.” *Political Psychology*. <https://doi.org/10.1111/pops.12663>.
- Bishop, Michael A., and J. D. Trout. 2004. *Epistemology and the Psychology of Human Judgment*. New York: Oxford University Press.
- . 2008. “Strategic Reliabilism: A Naturalistic Approach to Epistemology.” *Philosophy Compass* 3 (5): 1049–65. <https://doi.org/10.1111/j.1747-9991.2008.00161.x>.
- Brandt, Mark J., and Willem W. A. Sleegers. 2021. “Evaluating Belief System Networks as a Theory of Political Belief System Dynamics.” *Personality and Social Psychology Review* 25 (2): 159–85. <https://doi.org/10.1177/1088868321993751>.
- Byrd, Nick. 2019. “What We Can (and Can’t) Infer about Implicit Bias from Debiasing Experiments.” *Synthese* 198 (2): 1427–55. <https://doi.org/10.1007/s11229-019-02128-6>.
- . 2022. “All Measures Are Not Created Equal: Reflection Test, Think Aloud, and Process Dissociation Protocols.” <https://researchgate.net/publication/344207716>.
- Byrd, Nick, and Paul Conway. 2019. “Not All Who Ponder Count Costs: Arithmetic Reflection Predicts Utilitarian Tendencies, but Logical Reflection Predicts Both Deontological and Utilitarian Tendencies.” *Cognition* 192 (November): 103995. <https://doi.org/10.1016/j.cognition.2019.06.007>.
- Byrd, Nick, Gabriela Gongora, Brianna Joseph, and Miroslav Sirota. 2022. “Tell Us What You Really Think: A Think Aloud Protocol Analysis of the Verbal Cognitive Reflection Test.” <https://researchgate.net/publication/341131477>.
- Calvillo, Dustin P., and Thomas J. Smelter. 2020. “An Initial Accuracy Focus Reduces the Effect of Prior Exposure on Perceived Accuracy of News Headlines.” *Cognitive Research: Principles and Implications* 5 (1): 55. <https://doi.org/10.1186/s41235-020-00257-y>.

- Carlson, Ryan W., Michel André Maréchal, Bastiaan Oud, Ernst Fehr, and Molly J. Crockett. 2020. “Motivated Misremembering of Selfish Decisions.” *Nature Communications* 11 (1): 1–11. <https://doi.org/10.1038/s41467-020-15602-4>.
- Carnap, Rudolf. 1950. “Empiricism, Semantics, and Ontology.” *Revue Internationale de Philosophie* 4 (11): 20–40.
- Connor, Paul, Emily Sullivan, Mark Alfano, and Nava Tintarev. 2020. “Motivated Numeracy and Active Reasoning in a Western European Sample.” *Behavioural Public Policy*, August, 1–23. <https://doi.org/10.1017/bpp.2020.32>.
- Davies, R., Ives, J., & Dunn, M. (2015). A systematic review of empirical bioethics methodologies. *BMC Medical Ethics*, 16(1), 15. <https://doi.org/10.1186/s12910-015-0010-3>
- De Cruz, Helen. 2020. “Believing to Belong: Addressing the Novice-Expert Problem in Polarized Scientific Communication.” *Social Epistemology* 34 (5): 440–52. <https://doi.org/10.1080/02691728.2020.1739778>.
- De Keersmaecker, Jonas, David Dunning, Gordon Pennycook, David G. Rand, Carmen Sanchez, Christian Unkelbach, and Arne Roets. 2019. “Investigating the Robustness of the Illusory Truth Effect Across Individual Differences in Cognitive Ability, Need for Cognitive Closure, and Cognitive Style.” *Personality and Social Psychology Bulletin*, June, 0146167219853844. <https://doi.org/10.1177/0146167219853844>.
- De Neys, Wim. 2006. “Dual Processing in Reasoning Two Systems but One Reasoner.” *Psychological Science* 17 (5): 428–33.
- , ed. 2018. *Dual Process Theory 2.0*. 1 edition. Routledge.
- Derreumaux, Yrian, Robin Bergh, and Brent Hughes. 2020. “Partisan-Motivated Sampling: Re-Examining Politically Motivated Reasoning Across the Information Processing Stream.” PsyArXiv. <https://doi.org/10.31234/osf.io/fb7e6>.
- Ditto, Peter H., Brittany S. Liu, Cory J. Clark, Sean P. Wojcik, Eric E. Chen, Rebecca H. Grady, Jared B. Celniker, and Joanne F. Zinger. 2018. “At Least Bias Is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives.” *Perspectives on Psychological Science*, May, 1745691617746796. <https://doi.org/10.1177/1745691617746796>.
- Doris, John M. 2015. *Talking to Our Selves: Reflection, Ignorance, and Agency*. OUP Oxford.
- Dorst, Kevin. forthcoming. “Rational Polarization.”
- Dreyfus, Hubert. 1986. *Mind Over Machine*. Oxford: Blackwell Publishers.
- Epstein, Seymour. 1994. “Integration of the Cognitive and the Psychodynamic Unconscious.” *American Psychologist* 49 (8): 709.
- Evans, Jonathan St. B. T. 2007. “On the Resolution of Conflict in Dual Process Theories of Reasoning.” *Thinking & Reasoning* 13 (4): 321–39. <https://doi.org/10.1080/13546780601008825>.
- Everett, Jim Albert Charlton, Cory J. Clark, Peter Meindl, Jamie B. Luguri, Brian D. Earp, Jesse Graham, Peter H. Ditto, and Azim F. Shariff. 2021. “Political Differences in Free Will Belief Are Associated with Differences in Moralization.” *Journal of Personality and Social Psychology* 120 (2): 461–83. <https://doi.org/10.1037/pspp0000286>.
- Ferrin, Asia. 2017. “Good Moral Judgment and Decision-Making Without Deliberation.” *The Southern Journal of Philosophy* 55 (1): 68–95. <https://doi.org/10.1111/sjp.12210>.
- Frederick, Shane. 2005. “Cognitive Reflection and Decision Making.” *Journal of Economic Perspectives* 19 (4): 25–42. <https://doi.org/10.1257/089533005775196732>.

- Funkhouser, Eric. 2020. "A Tribal Mind: Beliefs That Signal Group Identity or Commitment." *Mind & Language*. <https://doi.org/10.1111/mila.12326>.
- Gampa, Anup, Sean P. Wojcik, Matt Motyl, Brian A. Nosek, and Peter H. Ditto. 2019. "(Ideo)Logical Reasoning: Ideology Impairs Sound Reasoning." *Social Psychological and Personality Science* 10 (8): 1075–83. <https://doi.org/10.1177/1948550619829059>.
- Gershman, Samuel J. 2018. "How to Never Be Wrong." *Psychonomic Bulletin & Review*, May, 1–16. <https://doi.org/10.3758/s13423-018-1488-8>.
- Goodman, Nelson. 1983. *Fact, Fiction, and Forecast*. Harvard University Press.
- Groenendyk, Eric, and Yanna Krupnikov. 2020. "What Motivates Reasoning? A Theory of Goal-Dependent Political Evaluation." *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12562>.
- Hamilton, Alexander, James Madison, and John Jay. 1787. *The Federalist Papers*. Oxford University Press.
- Hannon, Michael. 2020. "Empathetic Understanding and Deliberative Democracy." *Philosophy and Phenomenological Research*. <https://doi.org/10.1111/phpr.12624>.
- Hornsey, Matthew J., Emily A. Harris, and Kelly S. Fielding. 2018. "Relationships among Conspiratorial Beliefs, Conservatism and Climate Scepticism across Nations." *Nature Climate Change*, May, 1. <https://doi.org/10.1038/s41558-018-0157-2>.
- Huffman, Steve. 2018. "In Response to Recent Reports about the Integrity of Reddit, I'd like to Share Our Thinking." *r/Announcements*. Reddit. 2018. [https://www.reddit.com/r/announcements/comments/827zqc/in\\_response\\_to\\_recent\\_reports\\_about\\_the\\_integrity/](https://www.reddit.com/r/announcements/comments/827zqc/in_response_to_recent_reports_about_the_integrity/).
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford University Press.
- Janis, Irving L., and Frederick Frick. 1943. "The Relationship between Attitudes toward Conclusions and Errors in Judging Logical Validity of Syllogisms." *Journal of Experimental Psychology* 33 (1): 73–77. <https://doi.org/10.1037/h0060675>.
- Johnson-Laird, P. N., and Marco Ragni. 2019. "Possibilities as the Foundation of Reasoning." *Cognition* 193 (December): 103950. <https://doi.org/10.1016/j.cognition.2019.04.019>.
- Kahan, Dan M. 2013. "Ideology, Motivated Reasoning, and Cognitive Reflection: An Experimental Study." *Judgment and Decision Making* 8: 407–24. [sjdm.org/journal/13/13313/jdm13313.pdf](http://sjdm.org/journal/13/13313/jdm13313.pdf).
- . 2016. "The Politically Motivated Reasoning Paradigm, Part 1: What Politically Motivated Reasoning Is and How to Measure It." In *Emerging Trends in the Social and Behavioral Sciences*, 1–16. American Cancer Society. <https://doi.org/10.1002/9781118900772.etrds0417>.
- Kahan, Dan M., and Jonathan C. Corbin. 2016. "A Note on the Perverse Effects of Actively Open-Minded Thinking on Climate-Change Polarization." *Research & Politics* 3 (4): 2053168016676705. <https://doi.org/10.1177/2053168016676705>.
- Kahan, Dan M., Asheley Landrum, Katie Carpenter, Laura Helft, and Kathleen Hall Jamieson. 2017. "Science Curiosity and Political Information Processing." *Political Psychology* 38 (S1): 179–99. <https://doi.org/10.1111/pops.12396>.
- Kahan, Dan M., Ellen Peters, Erica Cantrell Dawson, and Paul Slovic. 2017. "Motivated Numeracy and Enlightened Self-Government." *Behavioural Public Policy*, May, 54–86. <https://doi.org/10.1017/bpp.2016.2>.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.

- Kennett, Jeanette, and Cordelia Fine. 2009. "Will the Real Moral Judgment Please Stand Up?" *Ethical Theory and Moral Practice* 12 (1): 77–96. <https://doi.org/10.1007/s10677-008-9136-4>.
- Kinder, Donald R., and Nathan P. Kalmoe. 2017. *Neither Liberal Nor Conservative: Ideological Innocence in the American Public*. University of Chicago Press.
- Kitcher, Philip. 1990. "The Division of Cognitive Labor." *The Journal of Philosophy* 87 (1): 5–22.
- Klein, Gary A. 1998. *Sources of Power: How People Make Decisions*. MIT press.
- Koriat, Asher. 2019. "Confidence Judgments: The Monitoring of Object-Level and Same-Level Performance." *Metacognition and Learning*, July. <https://doi.org/10.1007/s11409-019-09195-7>.
- Kornblith, Hilary. 2012. *On Reflection*. OUP Oxford.
- . 2019. *Second Thoughts and the Epistemological Enterprise*. New York: Cambridge University Press.
- Korsgaard, Christine M. 1996. *The Sources of Normativity*. Cambridge University Press.
- Lee, Nathan, Brendan Nyhan, Jason Reifler, and D. J. Flynn. 2020. "More Accurate, But No Less Polarized: Comparing the Factual Beliefs of Government Officials and the Public." *British Journal of Political Science*. <https://doi.org/10.1017/S000712342000037X>.
- Levin, Margarita. 1978. "The Problem Of Knowledge In The Original Position," June. <https://doi.org/10.17161/AJP.1808.8917>.
- Levy, Neil. 2019. "Due Deference to Denialism: Explaining Ordinary People's Rejection of Established Scientific Findings." *Synthese* 196 (1): 313–27. <https://doi.org/10.1007/s11229-017-1477-x>.
- Littlejohn, Clayton. 2018. "The Right in the Good: A Defense of Teleological Non-Consequentialism in Epistemology." In *Epistemic Consequentialism*, edited by Kristoffer Ahlstrom-Vij Jeff Dunn. Oxford University Press.
- Lord, Charles G., Lee Ross, and Mark R. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology* 37 (11): 2098.
- Lynch, Michael Patrick. 2018. "Arrogance, Truth, and Public Discourse." *Episteme* 15 (3): 283–96. <https://doi.org/10.1017/epi.2018.23>.
- Mercier, Hugo, and Dan Sperber. 2017. *The Enigma of Reason*. Harvard University Press.
- Mele, Alfred R. 2008. *Free Will and Luck*. Oxford University Press.
- Nagel, Jennifer. 2012. "Intuitions and Experiments: A Defense of the Case Method in Epistemology." *Philosophy and Phenomenological Research* 85 (3): 495–527.
- . 2014. "Intuition, Reflection, and the Command of Knowledge." *Aristotelian Society Supplementary Volume* 88 (1): 219–41. <https://doi.org/10.1111/j.1467-8349.2014.00240.x>.
- Oyserman, Daphna, and Andrew Dawson. 2019. "Your Fake News, Our Facts: Identity-Based Motivation Shapes What We Believe, Share, and Accept." In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*, edited by Rainer Greifeneder, M. Jaffé, Eryn J. Newman, and Norbert Schwarz. London, UK: Psychology Press.
- Pennycook, Gordon, ed. 2018. *The New Reflectionism in Cognitive Psychology: Why Reason Matters*. 1 edition. Abingdon, Oxon ; New York, NY: Routledge.
- Pennycook, Gordon, Jonathan A. Fugelsang, and Derek J. Koehler. 2015. "What Makes Us Think? A Three-Stage Dual-Process Model of Analytic Engagement." *Cognitive Psychology* 80 (August): 34–72. <https://doi.org/10.1016/j.cogpsych.2015.05.001>.



- Pennycook, Gordon, and David G. Rand. 2019a. "Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking." *Journal of Personality*. <https://doi.org/10.1111/jopy.12476>.
- . 2019b. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality." *Proceedings of the National Academy of Sciences* 116 (7): 2521–26. <https://doi.org/10.1073/pnas.1806781116>.
- Persson, Emil, David Andersson, Lina Koppel, Daniel Västfjäll, and Gustav Tinghög. 2021. "A Preregistered Replication of Motivated Numeracy." *Cognition* 214 (September): 104768. <https://doi.org/10.1016/j.cognition.2021.104768>.
- Peters, Uwe. 2019. "Implicit Bias, Ideological Bias, and Epistemic Risks in Philosophy." *Mind & Language* 34 (3): 393–419. <https://doi.org/10.1111/mila.12194>.
- Pew Research Center. 2017. "Views on Global Warming and Environmental Regulation, Personal Environmentalism." October 5, 2017. [people-press.org/2017/10/05/7-global-warming-and-environmental-regulation-personal-environmentalism/](http://people-press.org/2017/10/05/7-global-warming-and-environmental-regulation-personal-environmentalism/).
- Raoelison, Matthieu, Esther Boissin, Grégoire Borst, and Wim De Neys. 2021. "From Slow to Fast Logic: The Development of Logical Intuitions." *Thinking & Reasoning*, April. <https://doi.org/10.1080/13546783.2021.1885488>.
- Rathje, Steve, Azim Shariff, and Simone Schnall. 2020. "Ideology Trumps Self-Interest: Continued Support for a Political Leader Despite Disappointing Tax Returns," April. <https://doi.org/10.31234/osf.io/bptqj>.
- Rawls, John. 1971. *A Theory of Justice*. Revised edition. Cambridge, Mass: Belknap Press.
- Rini, Regina. 2017. "Fake News and Partisan Epistemology." *Kennedy Institute of Ethics Journal* 27 (2): E-43-E-64. <https://doi.org/10.1353/ken.2017.0025>.
- Roets, Arne, Dries H. Bostyn, Jonas De Keersmaecker, Tessa Haesevoets, Jasper Van Assche, and Alain Van Hiel. 2020. "Utilitarianism in Minimal-Group Decision Making Is Less Common than Equality-Based Morality, Mostly Harm-Oriented, and Rarely Impartial." *Scientific Reports* 10 (1): 1–10. <https://doi.org/10.1038/s41598-020-70199-4>.
- Rubin, Emily B., Anna E. Buehler, Elizabeth Cooney, Nicole B. Gabler, Adjoa A. Mante, and Scott D. Halpern. 2019. "Intuitive vs Deliberative Approaches to Making Decisions About Life Support: A Randomized Clinical Trial." *JAMA Network Open* 2 (1): e187851. <https://doi.org/10.1001/jamanetworkopen.2018.7851>.
- Savulescu, Julian, Guy Kahane, and Christopher Gyngell. 2019. "From Public Preferences to Ethical Policy." *Nature Human Behaviour*, August, 1–3. <https://doi.org/10.1038/s41562-019-0711-6>.
- Schwenkler, John. 2018. "Self-Knowledge and Its Limits." *Journal of Moral Philosophy* 15 (1): 85–95. <https://doi.org/10.1163/17455243-01501005>.
- Shea, Nicholas, and Chris D. Frith. 2016. "Dual-Process Theories and Consciousness: The Case for 'Type Zero' Cognition." *Neuroscience of Consciousness* 2016 (1). <https://doi.org/10.1093/nc/niw005>.
- Sikimić, Vlasta, Tijana Nikitović, Miljan Vasić, and Vanja Subotić. 2020. "Do Political Attitudes Matter for Epistemic Decisions of Scientists?" *Review of Philosophy and Psychology*, August. <https://doi.org/10.1007/s13164-020-00504-7>.
- Sidgwick, Henry. 1874. *The Methods Of Ethics*. 7th ed. Chicago: Hackett Publishing.
- Sosa, Ernest. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge University Press.

- Stanovich, Keith E, and Richard F West. 1997. "Reasoning Independently of Prior Belief and Individual Differences in Actively Open-Minded Thinking." *Journal of Educational Psychology* 89 (2): 342.
- Stephens, Andreas, and Trond A. Tjøstheim. 2020. "The Cognitive Philosophy of Reflection." *Erkenntnis*, September. <https://doi.org/10.1007/s10670-020-00299-0>.
- Stich, Stephen P. 1990. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA, US: The MIT Press.
- Strohming, Nina. 2018. "Identity Is Essentially Moral." In *Atlas of Moral Psychology*, edited by Kurt Gray and Jesse Graham, 141–48. Guilford Publications.
- Szaszi, Barnabas, Aba Szollosi, Bence Palfi, and Balazs Aczel. 2017. "The Cognitive Reflection Test Revisited: Exploring the Ways Individuals Solve the Test." *Thinking & Reasoning* 23 (3): 207–34. <https://doi.org/10.1080/13546783.2017.1292954>.
- Talaifar, Sanaz, and William B. Swann. 2018. "Deep Alignment with Country Shrinks the Moral Gap Between Conservatives and Liberals." *Political Psychology*. <https://doi.org/10.1111/pops.12534>.
- Talisse, Robert B. 2019. *Overdoing Democracy: Why We Must Put Politics in Its Place*. New York, NY: Oxford University Press.
- Tappin, Ben M., Gordon Pennycook, and David G. Rand. 2021. "Rethinking the Link between Cognitive Sophistication and Politically Motivated Reasoning." *Journal of Experimental Psychology: General* 150 (6): 1095–1114. <https://doi.org/10.1037/xge0000974>.
- Thomson, Keela S., and Daniel M. Oppenheimer. 2016. "Investigating an Alternate Form of the Cognitive Reflection Test." *Judgment and Decision Making* 11 (1): 99–113. <https://psycnet.apa.org/record/2016-09222-009>.
- Van Bavel, Jay J., and Andrea Pereira. 2018. "The Partisan Brain: An Identity-Based Model of Political Belief." *Trends in Cognitive Sciences* 22 (3): 213–24. <https://doi.org/10.1016/j.tics.2018.01.004>.
- Vegetti, Federico, and Moreno Mancosu. 2020. "The Impact of Political Sophistication and Motivated Reasoning on Misinformation." *Political Communication* 0 (0): 1–18. <https://doi.org/10.1080/10584609.2020.1744778>.
- Velleman, J. David. 1989. *Practical Reflection*. Princeton University Press.
- Vonasch, Andrew James. 2016. "Cognitive Miserliness Preserves the Self-Regulatory Resource." Florida State University. <https://diginole.lib.fsu.edu/islandora/object/fsu%3A366130/>.
- Wallace, R. Jay. 2006. *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason*. Clarendon Press.
- Weisberg, D. S., Landrum, A. R., Hamilton, J., & Weisberg, M. (2020). Knowledge about the nature of science increases public acceptance of science regardless of identity factors. *Public Understanding of Science*, 0963662520977700. <https://doi.org/10.1177/0963662520977700>
- West, Emily A., and Shanto Iyengar. 2020. "Partisanship as a Social Identity: Implications for Polarization." *Political Behavior*, July. <https://doi.org/10.1007/s11109-020-09637-y>.
- Wheeler, Gregory. 2018. "Bounded Rationality." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2018. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/win2018/entries/bounded-rationality/>.