Cognitive Bias and the Constitution

Dan M. Kahan
Yale Law School

Follow this and additional works at: http://digitalcommons.law.yale.edu/fss_papers

Part of the Law Commons

Recommended Citation
http://digitalcommons.law.yale.edu/fss_papers/4693
COGNITIVE BIAS AND THE CONSTITUTION

DAN M. KAHAN*

INTRODUCTION

This article has reciprocal goals. The first is to use the study of public risk perceptions to add psychological realism to liberal constitutional theory. The second is to use this enriched understanding of constitutional theory to add normative depth to the study of public risk perceptions.

I will attempt to achieve these objectives by discussing two phenomena. The first, cultural cognition, is a psychological dynamic.¹ It comprises a set of related mechanisms that unconsciously motivate individuals to form perceptions of risk and related facts that cohere with important group commitments.

The second phenomenon, cognitive illiberalism, is a normative concept.² It describes a state of affairs in which enforcement of liberal political principles is defeated, not by willful, defiance but by unconscious bias.

My thesis is that cultural cognition is an important source of cognitive illiberalism. In the course of observing social norms and applying legal rules that implement liberal political principles, citizens and governmental decision-makers are prone to unconsciously impute harm and other socially undesirable consequences to behavior that denigrates their groups' cultural outlooks.

This tendency has largely evaded the attention of constitutional theorists. Their primary mission has been to develop conceptions of constitutional liberties—from equal protection of the law to free exercise of religion to freedom of speech—that assure the State does not condition legal obligations on conformity to a governmentally ap-

* Elizabeth K. Dollard Professor of Law and Professor of Psychology, Yale Law School.


proved political orthodoxy. The U.S. Supreme Court has embroidered a rich fabric of doctrine that attests to the influence and ingenuity of this distinctively liberal conception of constitutional democracy.

There has been essentially no attention, however, to the fit between these doctrines and the psychological capacities of those expected to obey and enforce them. As a result, constitutional law is effectively innocent of cognitive illiberalism, which subverts state neutrality, notwithstanding widespread commitment to it, among judges, jurors, legislators, and citizens generally. The nature of this threat and how it might be mitigated are lessons that theorists of constitutional law can learn from the study of risk perception, where the mechanisms of cultural cognition are more fully appreciated.

By the same token, those who study risk perception have something to learn from the antagonism between cultural cognition and the liberal underpinnings of constitutional theory. These commentators tend to engage cognition from one of two opposing normative orientations. One is relentlessly welfarist: the aim of public policy, on this view, is to secure a cost-effective equilibrium between risk regulation and collective health, safety, and prosperity; achieving this goal demands procedures that insulate the law from cognitive biases, which can distort perception of the dangers posed by putative risk sources and of the costs and benefits of policies to abate such threats. The alternative orientation is populist. Public evaluations diverge from the experts' expected-utility calculus, it holds, not because the public is irrational, but because the experts are morally obtuse. The public apprehends risk through value-laden emotions that steer them toward stances that rationally express their commitment to one or another visions of the good life and the just society.

I believe this debate, while important, reflects an egregiously truncated understanding of what is at stake, morally, in the politics of risk regulation. Just as constitutional theorists have failed to comprehend the vulnerability of liberal neutrality to cultural cognition, so the scholars of risk perception have neglected to recognize the pervasive

danger that cultural cognition poses, not merely to the efficient balancing of regulatory costs and benefits, and not merely to the responsiveness of risk regulation to public values, but also to the obligation of the State to refrain from predicking legal obligation and benefits on a political orthodoxy. As a result, those on both sides fail to even ask, much less answer, a key set of moral questions.

Or so I will try to show over the course of an argument that will progress in stages. Part I presents a model of cultural cognition as bias. Part II illustrates the model by showing how this bias can distort public perceptions of a novel risk source—nanotechnology. Part III shows how the same dynamics predictably frustrate the enforcement of a critical doctrine of constitutional law: the "non-communicative harm principle." Part IV links cultural cognition to the problem of cognitive illiberalism in constitutional law. Part V shows how the problem that cognitive illiberalism poses for constitutional law is endemic to regulation of risk generally. Part VI outlines a set of parallel "de-biasing" strategies to make both constitutional law and risk regulation more resistant to cognitive illiberalism. Then I will briefly conclude with a reflection on the deeper source of cognitive illiberalism in the conditions that enable a society to acquire an unprecedented volume of knowledge only by disabling it from certifying authoritatively what it knows.

I. CULTURAL COGNITION AS BIAS

I will define a "bias" as any cognitive dynamic that causes individuals to process information in a manner that systematically defeats their goals. I do not think cultural cognition is inevitably a bias by this definition. But it can be, and I will now try to be precise about how with the aid of a simple model.

It will help to start, though, with an even simpler model that illustrates how unbiased information processing might work. Imagine an individual is trying to make an accurate assessment of some proposition—say, that an accomplished athlete in her favorite sport used a performance-enhancing substance. Furnished evidence relating to that proposition, the fan trying to assess this charge should give the evidence the weight it is due, and revise her assessment of the likelihood of his having achieved success by these means accordingly.

Fig. 1. Unbiased information processing. An individual whose goal is to form an accurate assessment of some factual proposition revises her existing belief by giving due weight to new evidence. In Bayesian terms, she updates by multiplying her prior odds by the likelihood ratio (how much more consistent the evidence is with the proposition being true than with the proposition being not true) to form her revised odds.

Bayes' theorem supplies a logical formalization of this process (Fig. 1). In effect, it says a person should adjust upward or downward her estimation of the likelihood of a proposition in proportion to how much more consistent the new evidence is with that proposition being true than with it being not true. More precisely, to determine her revised estimation of the proposition expressed in odds, she should multiply her prior estimation expressed in odds by the likelihood ratio associated with the new evidence.7 Thus, in the example, if the athlete tested positive in a procedure that has a 95% accuracy rate, the fan should increase her previous assessment of the odds (say, 1:99 or 1%) that the athlete used a performance-enhancing substance 19-fold (to 19:99, or 16%), since a "true" positive is 19x (95:5) more likely than a "false" positive.8

8. If one thinks about, or better yet visualizes, such a problem as one involving the frequency of "true" and "false" positives within a large sample of cases, the logic of Bayes' Theorem—it is just a logical truism, albeit a surprising and strikingly beautiful one—becomes easier to grasp. See David Spiegelhalter, Mike Pearson & Ian Short, Visualizing Uncertainty About the Future, 333 Sci. 1393, 1394 fig. 2 (2011). Indeed, it even becomes intuitive after only a modest number of exercis-
The numbers are not really so important here. What matters is that an unbiased information processor *revises her existing beliefs* (her "priors") consistent with the *weight* ("likelihood ratio") properly assigned to any new information she receives.

![Diagram](image)

**Fig. 2. Confirmation bias.** An individual displays confirmation bias when she determines the weight to be assigned the evidence by its consistency with her existing position. In Bayesian terms, the likelihood ratio is endogenous to the individual's priors. Depending on how strong the influence of her priors is on her weighting of the evidence, she will not revise a mistaken belief or will not revise it as quickly or dramatically as she should when exposed to valid contrary evidence.

That will not happen if her thinking displays "confirmation bias." (Fig. 2) This is the situation in which a person selectively credits or discounts evidence in patterns that reinforce what she already believes. Upon learning the athlete tested "positive," the confused fan initially exclaims, "Yes, of course! I told you it was virtually inconceivable that he would cheat—the positive test result confirms his innocence!" Advised that a "positive" result actually signifies use of the banned substance, she now sourly responds, "Pfft, that means nothing:

See Peter Sedlmeier & Gerd Gigerenzer, *Teaching Bayesian Reasoning in Less Than Two Hours*, 130 J. EXPERIMENTAL PSYCHOL. 380 (2001). Say we started with a group of 2,000 athletes, knowing only twenty (or 1%) had used performance enhancing drugs. Assuming a test with a 95% accuracy rate, we would expect ninety-nine of the 1,980 "clean" athletes to register false positives (1,980 x 0.05). We would expect nineteen of the twenty "dirty" athletes (20 x 0.95) to register true positives (and thus one to register a false negative). Accordingly, of the 118 athletes who tested positive, we would believe only 16% (19+108) to be guilty of using performance-enhancing drugs. If we knew nothing more about who was genuinely clean and dirty, we would thus estimate there was a 16% probability that any particular athlete who tested positive had in fact used drugs.
such tests frequently turn out to be mistaken." Here, the fan is not revising her beliefs based on the weight of the new evidence; rather she is using her pre-existing beliefs to determine how much weight the evidence should be assigned. In Bayesian terms, the likelihood ratio has become entangled with (is endogenous to) her priors. As a result, she will persist in a mistaken belief for much longer than she should—maybe even forever, depending on the degree of entanglement—despite being continuously exposed to accurate corrective information.9

"Cultural cognition as bias" is like this but worse. In the next model—which is the model I mean to focus attention on in this Part—individuals are posited to have "cultural predispositions." For now, we can just think of these as tendencies that members of a group share to find some claims about risk (or related types of facts) more convincing than others. In Bayesian terms, cultural predispositions can be viewed as the source of individuals' "priors," their pre-existing beliefs. But cultural predispositions will also affect the weight they assign to new evidence—their determination of the "likelihood ratio," in Bayesian terms. That is, they will tend to credit or discount new information depending on whether it supports a conclusion compatible with their predispositions (Fig. 3, panel A).

![Diagram](attachment:diagram.png)

**Fig. 3. Cultural cognition as bias.** As a result of cultural cognition, individuals selectively credit and discount evidence in patterns that reinforce what those who share their cultural commitments are predisposed to believe (A). They will (or will often) appear (B) to be displaying confirmation bias, because the same phenomenon—

cultural predisposition—that is causing their existing beliefs (their "priors" in Bayesian terms) is determining the weight ("likelihood ratio") they assign new evidence. Under these circumstances, though, individuals will not only resist revision of their beliefs but will also persist in a state of disagreement with other individuals who hold alternative cultural predispositions.

In this model, we will often see what looks like confirmation bias. There will often be—or at least appear to be—that dynamic's signature entanglement between priors and likelihood ratio. That will be the case because the same thing—cultural predispositions—is causing both individuals' existing beliefs and the weight that they assign new evidence (Fig. 3).

But because cultural predispositions vary across groups, there will also be persistent divisions of opinion. The members of groups with diverse predispositions will start off with conflicting priors, and when they are furnished with new evidence, they will assign opposing degrees of significance to it depending on its consistency with what they are already predisposed to believe. Despite their common exposure to pertinent evidence, they will not converge on the most likely assessment of some proposition; rather, because they are giving differential effects to the weight of that information, they are likely to grow even farther apart.

If their goal is to form accurate beliefs, this mode of information processing will systematically frustrate their ends. And if what is at stake in getting things right is not something prosaic, like whether the star of one or the other cultural group's favorite sports team is a cheat, but instead something important—like what policy their elected representatives should adopt to secure their common welfare, or whether some form of conduct of extreme moral significance to one group should be prohibited because it is exposing members of the other to genuine risks of harm—then the consequences of being unable to converge on the best understanding of the facts might be very bad indeed.

II. CULTURAL COGNITION AND NANOTECHNOLOGY RISK PERCEPTIONS

Nanotechnology comprises a set of engineering techniques for exploiting the quantum-mechanical properties of atomic-scale particles. Most members of the public do not know this; in fact, they have either never heard of nanotechnology or have heard something but are not really sure what it is.
My collaborators in the Cultural Cognition Project and I were curious how individuals' perceptions of the risks and benefits of nanotechnology were likely to take shape as they learned more about it. In particular, we were eager to study the formation of nanotechnology risk perceptions to test certain hypothesis relating to cultural cognition. So, we conducted an experimental study involving a large, nationally representative sample of U.S. adults.10

The first thing we did was measure our study subjects' "cultural worldviews." We did so with attitudinal scales patterned on a framework associated with the late anthropologist Mary Douglas. Douglas's scheme characterizes people's worldviews—their preferences, essentially, about how society or other collective endeavors should be organized—along two cross-cutting dimensions: "Individualism-communitarianism" ("IC") and "Hierarchy-egalitarianism" ("HE") (Fig. 4).11

The distinguishing characteristics of these outlooks are fairly intuitive and are in fact featured in a variety of comparable schemes found in the social sciences. Basically, the more inclined a person is to feel that individuals should be responsible for securing the conditions of their own flourishing without assistance or interference from the State or some comparable collective entity, the more "individualistic" she is; the more inclined she is to feel that securing the conditions of individual flourishing is the duty of the collective, and that the collective should be empowered to override individual decisions that compromise this duty, the more "communitarian" she is. "Hierarchically" inclined individuals believe that authority within societies and other collective entities should be apportioned on the basis of stratified social roles and offices that are conspicuous and largely fixed. People who believe that who can tell whom to do what should not depend on stratified distinctions of this sort are "egalitarian" in disposition.

The theory of cultural cognition posits that individuals' perceptions of risk and related facts are motivated by dispositions that cohere with shared "worldviews," or understandings of how society or other collective enterprises should be organized. The worldviews consist of combinations of values formed by the intersection of two cross-cutting continua, "Hierarchy-egalitarianism" and "Individualism-communitarianism."

According to Douglas and collaborator Aaron Wildavsky, we should expect individuals to form risk perceptions that reflect and reinforce their commitment to one or another combination of values associated with the HE-IC framework. "Hierarchical individualists," they surmised would be skeptical of environmental risk claims, the widespread acceptance of which would justify restrictions on commerce, industry, and markets—forms of activity important not only materially but symbolically and emotionally in the way of life of such persons. "Egalitarian communitarians," in contrast, tend to be morally suspicious of commerce, industry, and markets, to which they attribute various forms of social inequity. They would thus be receptive to claims that such activity is dangerous and should be restricted in the interest of public safety and health.

Douglas and Wildavsky offered only impressionistic (bordering on polemical) support for their claims when presenting them in their most famous work, *Risk and Culture* (1982). But subsequent observational studies, including ones conducted by Karl Dake, one of Wildavsky's graduate students, generated evidence supportive of their conjectures. Douglas and Wildavsky were focused mainly on disputes over nuclear power, but studies suggest the predispositions they attributed to hierarchical individualists and egalitarian communitarians generalize to other environmental and technological risks, including ones associated today with climate change. Studies have tested and confirmed hypotheses relating the HE-IC scheme and various other types of risks, too, including ones that more sharply divide “hierarchical communitarians” and “egalitarian individualists.”

Nanotechnology, though, does not cut as large a “cultural profile”—as it were—as climate change, nuclear power, or many of the other risk disputes, such as ones over gun possession and gun control, that had been featured in tests of the Douglas-Wildavsky theory. Accordingly, we saw nanotechnology as an appropriate focus for examining the role of cultural predispositions in the formation of risk perceptions.

Moreover, because we could be confident that most subjects would not have informed beliefs prior to the study, their reactions to information struck us as more amenable to experimental manipulations aimed at identifying the psychological mechanisms that link cultural worldviews to risk perceptions. This was a matter on which Douglas and Wildavsky had little convincing to say; indeed, they were most persuasive when they simply ignored the question. On the few occasions in which they addressed mechanisms of belief formation they resorted to functionalist styles of explanation that had been devastatingly criticized and appropriately abandoned by thoughtful social scientists decades earlier. We sought to “psychologize” the cultural theory of risk: our hypothesis was that individuals unfamiliar with nanotechnology would tend to make sense of it in patterns that reflect-

ed their cultural predispositions toward other forms of environmental and technological risk.

That was what we found (Fig. 5). In our study, we solicited subjects' reactions to the risks and benefits of nanotechnology after dividing them into two experimental treatment groups. The members of one—the "no information" group—received no information about nanotechnology except for a spare definition: a technology for manipulating particles on the scale of atoms that have special properties as a result of their small size. The members of the other got "balanced information": a paragraph describing some of the potential benefits people expected of nanotechnology (its use for manufacture of better sports equipment and cosmetics; its possible use for environmental cleanup) and another describing some of the potential risks people were concerned about (damage to health from ingestion; use to enable government monitoring). In the no-information group, most subjects indicated that they expected benefits to predominate over risks but there was no cultural variation—or variation of any systematic kind—to speak of; there was mainly noise, as one would expect among a group of people who really had no idea what they were offering an opinion on. But in the "information exposed" group, subjects with hierarchical individualistic outlooks, on the one hand, and those with egalitarian communitarian outlooks, on the other, divided sharply, the former seeing benefits predominating over risks and the latter risks predominating over benefits.

![Fig. 5. Impact of information exposure on nanotechnology risk-benefit perceptions. Relative to subjects in the "no information" condition, those in the "information-exposed" condition polarized](image-url)
along lines consistent with their cultural predispositions toward other environmental and technological risks.16

This is the dynamic reflected in the "cultural cognition as bias" model (Fig. 3, panel A). The subjects in the "information exposed" group all received the same mix of information. But the effect they gave to it—which parts they credited and which they discounted—depended on their predispositions toward environmental and technological risks generally. As a result, they polarized relative to their counterparts who lacked such information.

In this particular case, moreover, it is clear that the effect is not really a result of confirmation bias. Knowing next to nothing about nanotechnology, the subjects in the information-exposed condition had no particularly meaningful priors, or at least none that they were not ready to revise immediately upon being furnished with new information. How they formed (or revised) their views depended not on what they believed about nanotechnology before the study, but on what they were culturally predisposed to believe about a certain class of risks. Because we can see that the weight the subjects assigned to the evidence was motivated by their predispositions and not by their priors, it would be more accurate to call this "predisposition bias." Of course, if predisposition bias persists, such individuals will thereafter selectively credit or discount evidence in patterns that reinforce the beliefs they formed in response to their initial exposure to information. At that point, we will observe confirmation bias's signature entanglement between priors and likelihood ratio (or weight)—although in fact we will know that the correlation between the two is spurious, in fact a third influence, cultural predispositions, is causing both the individuals' priors and their biased weighting of new evidence (Fig. 3, panel B).

III. CULTURAL COGNITION AND THE NON-COMMUNICATIVE HARM PRINCIPLE

I now turn to a legal example—a constitutional law one—that fits the cultural cognition model. The example involves the so-called "non-communicative harm" principle.

The study was in the form of a mock jury experiment. In it, we instructed subjects (approximately 200 adults drawn from a nationally representative panel) to imagine they were the jurors in a case in

which a group of political protestors was suing the police for forcing the protestors to halt a demonstration. The police maintain that they broke up the protest because the protestors were physically intimidating passersby and obstructing access to a public building. The protestors deny this claim; they were engaged in passionate but non-coercive advocacy, they say, and were ordered to stop only because their message offended members of the public and the police.17

Fortunately, there is a videotape of the demonstration (Fig. 6), we advise the subjects. Both sides agree that it accurately depicts the behavior of the protestors. Both in fact assert that the video furnishes decisive evidence for their side of the case. Please view the tape, the subjects are instructed, and determine whose story—the protestors or the police—it most supports.

![Fig. 5. Experiment video. Subjects viewed a video of a political protest to determine the accuracy of competing characterizations of the behavior of the demonstrators as involved in intimidation or persuasion.](image)

With this design, we intended to track certain elements of a classic 1950s psychology study reported in an article entitled They Saw a Game.18 The subjects in this study were students from two Ivy League colleges in the U.S. The researchers instructed them to watch a film of a


football game between their two schools and to evaluate a set of controversial calls by the referees. The student subjects from Dartmouth, they found, believed the penalties assigned to Princeton were perfectly correct but that the ones called on Dartmouth were clearly in error; the student subjects from Princeton reported that the referees obviously got it wrong in the calls they made against Princeton but right in the calls they made against Dartmouth. The emotional stake the students had in experiencing solidarity with their institutions, the researchers concluded, had unconsciously shaped what they saw in the film. We hypothesized that a similar motivation to form perceptions supportive of one’s cultural commitments could similarly bias perceptions of individuals in the role of factfinders in a legal proceeding.

**Abortion Clinic Condition**

Outside the reproductive health clinic 15 minutes before it was scheduled to open.

Scene at the entrance of clinic. Patient approaches but does not enter.

**Recruitment Center Condition**

Outside the campus recruitment center 15 minutes before it was scheduled to open.

Scene at the entrance of recruitment center. Student approaches but does not enter.

Fig. 6. Experimental manipulation. Half the subjects were advised that the protest was critical of abortion rights and occurred outside and abortion clinic. The other half were advised the protest took place outside a campus recruitment center on a day in which the military was interviewing students interested in enlisting; the aim of the protest (those subjects were told) was to register objection to the military’s policy of excluding openly gay and lesbian individuals from service.

Unlike the study in *They Saw a Game*, our study—which we described in a paper entitled *They Saw a Protest*—had an experimental manipulation. We told half the subjects that the demonstration had occurred outside an abortion clinic, and that the protestors were criticizing the right of women to terminate unwanted pregnancies (Fig. 7).
We told the other half, in contrast, that the demonstration had occurred outside a college career-placement office on a day when the U.S. military was interviewing students interested in enlisting; the protestors, we advised the subjects, were objecting to the military's then-existing "Don't Ask, Don't Tell" policy, which barred military service by openly gay and lesbian individuals. The wording of the signs of the protestors, who in fact were demonstrating against neither abortion rights nor "Don't Ask, Don't Tell," were blurred out—at the direction of the court, we advised the subjects, in order to assure that jurors were not unconsciously influenced by the protestors' messages, which, we informed them, should play no role in their determination of the case.

What was supposed to determine the outcome, we told the subjects, was the conformity of the parties' conduct (the protestors' and the police's) to the law. The wording of the operative legal statute was also manipulated. In the "abortion clinic condition," subject were supplied with a law that made it illegal for any individual to "intentionally interfere with, obstruct, intimidate, or threaten any person . . . seeking to enter, exit, or remain lawfully on premises of any hospital or medical clinic . . . licensed to perform abortions"; in the "military recruitment center condition," the law prohibited any individual from directing the same forms of coercive behavior at any person "seeking to enter, exit, or remain lawfully on premises of any facility in which the U.S. military is engaged in recruitment activity." Both laws also authorized the police to "order" any person in violation of the statute "to desist and to leave the immediate vicinity."

After the subjects watched the video, we asked them to tell us what they saw. Subjects in both conditions indicated agreement or disagreement with a series of statements relating to the behavior displayed in the video. True or false—the protestors "blocked" entry into the building (either the abortion clinic or military recruitment center, depending on the experimental condition) and "screamed in the faces" and "shoved" people trying to enter the facility? "People trying to enter asked the police to help them." Or alternatively, potential entrants "did not feel threatened or intimidated," but rather steered clear of the building because they "just didn't want to have to listen" to the protestors' message. The reason the police broke up the protest is not that the "protestors were interfering with, intimidating, obstructing, or threatening anyone," but rather that members of the public and the police found the protestors "annoying."
These (and like) facts were selected to focus attention on conduct relevant both to applying the controlling statute and to assessing the constitutionality of the police order to halt the protest. Whether government officials can disperse protestors in a case like the one featured in our study depends on the interest or goal they are trying to attain. They cannot restrict such activity because they disagree with the protestors' position, or because members of the public are offended, angered, or distressed (even to the point of disorderly retaliation) by the ideas the protestors are advocating. Under the First Amendment, the government simply cannot "count as a harm" subject to legal preemption or redress any state of affairs "that grows out of ... the way people can be expected to react to [a speaker's] message." Rather, it must confine its attention to the "non-communicative impact" of expressive activity—consequences the undesirability of which can be identified independent of anyone's negative reaction to an idea. Thus, it can legitimately prohibit protestors from blocking, obstructing, intimidating, and threatening others—by physically "shoving" them, say, or by "screaming in their faces"—because the harms such behavior imposes do not depend on the message the protestors are expressing or indeed whether they are trying to convey any idea whatsoever.

This is the "non-communicative harm principle," a central pillar of contemporary First Amendment doctrine. The versions of the hypothetical statute we directed our subjects to apply—like the real-world laws they were patterned on—were specifically designed to satisfy it.

Indeed, among the real-world cases in which the non-communicative-harm principle has been applied, is one that involved the very protestors whose behavior was featured in our study video. In reality, those individuals were neither pro-life nor gay-rights activists. Rather, they were members of the Westboro Baptist Church, an anti-gay hate group. In Snyder v. Phelps,22 the U.S. Supreme Court overturned a $5 million judgment for "emotional distress" entered against the Church for picketing the funeral of a soldier, whose death in Iraq the Church members identified as an instance of divine retribution for societal tolerance of homosexuality. The Court did not question the upsetting consequences of the Church members' behavior for members

21. See id.
of the soldiers’ family, but observed that “any distress occasioned by Westboro’s picketing turned on the content and viewpoint of the message conveyed”—no one would have been upset if the Church members had been expressing gratitude for the fallen soldier’s sacrifice—“rather than any interference with the funeral itself.” Our study in effect anticipates the next case, in which the Church members sue to challenge enforcement of a statute that prohibits them from “interfering with the funeral itself”—by “obstructing, intimidating, or threatening” those in attendance. There will not be any dispute about whether the State can prevent (or punish) the Church members from engaging in such conduct; but there will, inevitably, be one about whether in fact they did—even if the events are captured on video.

The point, again, of the study was to assess whether the ability of the jurors to distinguish between constitutionally non-cognizable “communicative harms” and legally proscribable “non-communicative” ones would be influenced by cultural cognition. We again measured the cultural outlooks of our subjects using scales derived from Mary Douglass’s HE-IC framework. The cultural worldviews associated with the framework—“hierarchical individualist,” “hierarchical communitarian,” “egalitarian individualist,” and “egalitarian communitarian”—were for the subjects in our study what the school affiliations were for the subjects in the They Saw a Game study. That is, consistent with cultural cognition, the stake the subjects had in affirming their commitments to the cultural groups is what we expected to unconsciously shape their perceptions of the behavior displayed in the video.

Again, this is what we found. In effect, individuals who held opposing cultural outlooks disagreed with one another about what they saw on the tape—blocking and shoving protestors or annoyed and ashamed onlookers; “screaming in the face” or impassioned but non-coercive advocacy—when they were in the same experimental condition and thus had the same impression about the message the protestors were expressing. However, subjects who had the same cultural outlooks but who were assigned to different experimental conditions—and who thus had different impressions of the message the protestors were expressing—disagreed with one another about what they had seen (Fig. 8). The direction and magnitude of these effects (which were most dramatic for hierarchical communitarians and egalitarian individualists) were consistent with hypotheses we had formed.

23. Id. at 1219.
based on the observed predisposition of individuals with these outlooks toward risk claims that variously affirmed and threatened norms integral to the group's competing understandings of the status conferring social roles appropriate for men compared to women.

Fig. 8. Cultural variance in fact perceptions across experimental conditions. In the experiment, what subjects of diverse cultural outlooks saw on the video tape depended on which experimental condition—abortion clinic military-recruitment facility—they were assigned to.

IV. COGNITIVE ILLIBERALISM IN CONSTITUTIONAL LAW

A cognitive dynamic should be seen as a bias, I stated in Part I, when it systematically defeats the ends an actor is trying to achieve by engaging with information. I now want to identify the ends that cultural cognition undermines, and how, in real-world settings akin to those in the two experiments.
In Part I, I outlined how cultural cognition can be a bias for someone whose goal is to form an accurate assessment of some proposition. This is indeed a part of the problem here. But by reformulating this point now in relation to the *They Saw a Protest* study, the way in which the resulting errors threaten principles of liberal democracy—both in constitutional law and in the regulation of risk more generally—should become readily apparent.

In that experiment, individuals furnished with the same evidence formed highly divergent conclusions about what happened. This result by itself, however, does not support any inference of biased information processing (Fig. 9).

Fig. 9. Polarization despite unbiased information processing in political protest case. Decision-makers who start with different priors (about, say, the incidence of intimidating behavior by abortion or gay-rights protestors) and who give the same effect to evidence (a video tape of a protestor screaming in the face of a person trying to enter a facility) might still fail to converge on a perception of some disputed fact even when they process the information in an unbiased fashion.

Coming into a trial like the one featured in *They Saw a Protest*, prospective factfinders will have diverse existing beliefs—priors—about the frequency of intimidating conduct by protestors of different sorts. The priors of people with different cultural outlooks are likely to vary systematically—not necessarily because they are biased but because their cultural commitments are likely to correlate with different
life experiences and different information sources, the materials from which one might expect them to have formed their expectations about how abortion protestors or gay rights activists behave.

Jurors who start with different priors might continue to disagree about the facts even if they all agree about the significance of the evidence presented at trial. Imagine jurors watch a video of a protestors "screaming in the face" of a person trying to enter a military recruitment facility. All might agree—after all, this hardly seems to be a matter about which Hierarchical Communitarians and Egalitarian Individualists would be divided—that evidence of a demonstrator screaming in the face of an onlooker, while not in itself conclusive, makes the inference that the protestors were engaged in "intimidation" rather than "persuasion" more compelling (perhaps the jurors will surmise that roughly 50% of the protests in which "screaming in the face" is observed, but only 5% of ones in which it is not, cross the "persuasion"-"intimidation" line; likelihood ratio = 10). In that case, they will all be more inclined to believe that there was intimidation in the protest in question than they would have been without seeing the video. But if Egalitarian Individualists started with a much lower estimate of the frequency of intimidation in such events in general (1:100) than did the Hierarchical Communitarian (1:4), they would still end up disagreeing about what happened in this particular case (Egalitarian Individualists: 1:10 × 10 = 1:10 or 9% likelihood of intimidation; Hierarchical Communitarian 1:4 × 10 = 5:2 or 71%). That is, they will disagree despite everyone having assessed the evidence consistently with the model of "unbiased information processing" that I outlined at the very beginning of Part II (Fig. 1).
Confirmation bias in political protest case. If decision-makers who start with different priors (say, on incidence of "intimidation" in a particular form of protest) are unconsciously motivated to assign to the same evidence (a video of a protestor engaging an onlooker trying to enter a facility) weight that reinforces their priors, then they are displaying confirmation bias. In the face of common evidence, they will become even more polarized.

That is not what happened, though, in the *They Saw a Protest* study. There, the subjects did not agree about what they saw on the video. In the military-recruitment center condition, most Hierarchical Communitarians observed a protestor screaming in the face of someone trying to enter the facility (likelihood ratio $= 10$); yet most Egalitarian Individualists did not—they saw a protestor who was engaged only in impassioned advocacy of the sort perfectly consistent with persuasion (likelihood ratio $= 1$). What is more, the proportions of Hierarchical Communitarians and Egalitarian Individualists who saw one thing or another were reversed in the abortion-clinic condition. Let us assume (for a moment) that Hierarchical Communitarians and Egalitarian Individualists started with opposing priors on the frequency of intimidation in both military-recruitment protests ($1:4$ and $1:100$, respectively) and abortion-clinic protests ($1:100$ and $1:4$, respectively). In that case, it would be clear that the subjects were weighting the video evidence based on their preexisting beliefs (Fig. 10). In other words, we would be observing the entanglement of "likelihood ratio" with "priors" that is the signature of confirmation bias (Fig. 2).

But in fact, we do not have any information about the subjects' priors. We only know what their cultural predispositions were. The
results of the experiment, then, show that they, like the subjects in the nanotechnology experiment, were displaying a form of predisposition bias:24 that is, they were attributing to the evidence the weight most congenial to their cultural worldviews ("Look: she's screaming in someone's face!" → likelihood ratio = 10; "C'mon! There's no intimidation here—just singing and chanting of the sort the 1960s Civil Rights protestors engaged in!" → likelihood ratio = 1) (Fig. 11).

Fig. 11. Cultural cognition as bias in political protest case. If decision-makers with diverse cultural predispositions are unconsciously motivated to assign to the same evidence (a video of a protestors engaging an onlooker trying to enter a facility) weight that is congenial to their cultural predispositions, then they are displaying cultural cognition as bias. In the face of common evidence, they will become polarized, whether or not they were before they were exposed to the evidence.

It is likely that people with values of the sort we measured do have opposing priors on the relative frequency of intimidation at abortion-clinic and military-recruitment center protests. But given how few of them are likely to have had any first-hand experience with either sort of demonstration, the effect of predisposition bias on their filtering of information about such activities is the most likely reason why (cf. Fig. 3, panel A).

24. See supra pp. 375-76.
The normative ends systematically defeated by this pattern of information processing involve something much more fundamental than "accurate estimations of factual propositions" in the abstract. The whole point of the non-communicative harm principle is to enforce a stance of liberal neutrality in law. In order to avoid picking sides in disputes over the best way of life, the State cannot treat public aversion to a disfavored view as a basis for restricting advocacy on behalf of it; it can impose impositions on those whose profession of unorthodox values provoke public resentment only to avert harms (the infliction of bodily injury, the interference with freedom of motion, the imposition of reason-disabling states of fear, and the like) that can be defined (experienced) independently of a commitment to any particular vision of how to live. But if as a result of cultural cognition, those charged with enforcing this doctrine impute non-communicative harms more readily to behavior that denigrates their worldview than to behavior that affirms it, the law will end up conditioning the right to participate in political protests on the speakers' conformity to privileged cultural values in exactly the way it would have if the State were free to treat denigration of those commitments as a legally cognizable harm. Biased factfinding begets a culturally biased regime of law.25

This is how cultural cognition generates "cognitive illiberalism." Legislators and law enforcers need not be consciously motivated to condition obligations and entitlements on adherence to dominant norms. Nor need courts falter in their dedication to impartially assessing whether an illiberal ambition of this sort has corrupted the political process. They can all be genuinely committed to abiding by the injunction that the State refrain from imposing a sectarian cultural orthodoxy and confine law to the promotion of secular ends—ones the enjoyment of which are open on equal terms to persons of all cultural persuasions. Yet, despite their intentions they could all end up participating in the culturally partisan allocation of legal benefits and burdens because of their tendency to perceive behavior that threatens their values as threatening their collective health, safety, security, or prosperity.

The occasions for this sort of unconscious subversion of constitutional doctrine, moreover, are not limited to protection of free speech. Myriad doctrines assimilate constitutional freedoms to the liberal ideal of neutrality through devices analytically equivalent to the non-

communicative harm principle. Respecting the free exercise of religion requires the State to demonstrate that a law that "incidentally burdens" religious practice is both "generally applicable" and sincerely motivated by "secular" ends. The imposition of legal disadvantage on a particular group violates equal protection unless it "can be explained by reference to legitimate public policies" unrelated to simple "animosity toward" the group's unorthodox values. The right to personal liberty secured by due process protects individuals' "decisions... concerning the intimacies of their physical relationship"; the desire to "enforce... on the whole society" standards of "private conduct" that originate in "religious" or equivalent "conceptions of right and acceptable behavior" is not a "legitimate government purpose" for interfering with such decisions absent a showing that they are otherwise "harmful to others." All of these doctrines, too, presuppose that those responsible for complying with and enforcing the Constitution have the psychological capacity to make unbiased assessments of the consequences of behavior (and laws) that selectively threaten and affirm their cultural worldviews.

If the study of cultural cognition of risk teaches us anything, it is that this capacity cannot be taken for granted. But in effect, that is what constitutional theorists have done. They (or the ones whose work I am interested in, in any case) have set for themselves the task of explicating liberal conceptions of the individual liberties provisions of the Constitution, and fashioning decision-making frameworks attuned to securing the form of neutrality that liberalism envisions. But the theorists have had had virtually nothing (or nothing empirically informed) to say about the cognitive feasibility of the neutral assessments of information that these doctrines depend on.

I have been addressing cognitive illiberalism as if the problem had to do entirely with the influence of cultural cognition on those applying constitutional doctrines. I do indeed believe the dynamics observed in They Saw a Protest furnish grounds for believing that real-world juries experience similar distortions in perception. I think it is plausible to assume, too, that judges, in exercising their official functions, will not be immune from the impact of cultural cognition, although I do not think a study based on a general population sample adds any weight to that surmise; judicial habits of mind reflect a distinctive and intense

experience of professionalization that treat evidence of how non-judges (and in particular non-lawyers) reason as evidence of how judges do. But, in fact, the difficulty that cognitive illiberalism poses for constitutional law does not depend on the impact of cultural cognition on either juries or judges.

A study like They Saw a Protest is best understood as modeling, and testing hypotheses about, the manner in which members of the public react to decisions informed by doctrines like the non-communicative harm principle. Furnished with even less evidence than were the study subjects, ordinary citizens nevertheless form strong convictions about the facts in cases like the one in our experiment. This is something familiar to all of us (particularly any of us who have observed what happens in law school classrooms when controversial cases are taught and discussed). The study results furnish corroboration that the reason people do form strong convictions is cultural cognition. Individuals are motivated to form assessments of information—not just empirical study results, logical arguments, and indicia of source credibility but even brute sense impressions—that reinforce their group commitments. The defining groups of consequence in this setting, moreover, are the ones that are featured in the cultural theory of risk.

Research suggests that identity-protective forms of motivated cognition are marked by a characteristic feedback pattern. We are generally aware of the tendency of those who harbor group commitments different from ours to form assessments of evidence supportive of their own groups' position on contested facts; we are less likely to recognize that same tendency in ourselves and others who share our own group commitments, at least during moments when it is actually affecting them. This asymmetry—known as "naive realism"—can breed acrimony in inter-group deliberations. Each groups' members discern—and decry—bad faith in the other's evidently self-serving assessments of the weight of the evidence. The resulting cycles of denunciation and counter-denunciation furnish additional evidence,


oppositely weighted by members of each group, that their side’s position is the most defensible one.31

Indeed, in the resulting climate of polarization, opposing factual claims become symbols of the moral competence and status of competing cultural groups. As statistics-pervaded as it appears to be, the dispute over whether “more guns” mean “more crime” or “less” is really about what sort of society we want to live in—one in which individuals retain the responsibility and power to secure the conditions of their own well being, or instead submit to a condition of servile dependence on the state for their protection;32 one “based on an internal . . . balance of terror,” in which “each relies on himself for survival,”33 or one in which the responsibility to protect all individuals from violence is collectively shared through democratically accountable agencies of law enforcement and through reciprocal norms of trust and concern.34 The debate over whether human activity is heating the planet is “in reality a struggle for the soul of America”:35 between those who believe relentless consumption evinces an outlook of base and heedless self-

31. Kahan, supra note 2, at 130-42.
32. See, e.g., Daniel White, Why People Carry Guns, EXAMINER.COM (Mar. 16, 2006), http://www.examiner.com/article/why-people-carry-guns ("Taking accountab[ility] for your own safety and well being is something that used to be expected in this country. Now, it is more common for that obligation to be passed along to someone else, whether the police, the government, or a higher power. However, there are still plenty of people who still choose to be responsible for protecting themselves and their loved ones themselves."); Robert J. Cottrol, Gun Control Poses a Threat to Self-Defense, In HOW CAN GUN VIOLENCE BE REDUCED? 22, 26 (Laura K. Egendorf ed., 2002) ("A people incapable of protecting themselves will lose their rights as a free people, becoming either servile dependents of the state or of the criminal predators who are their de facto masters.").
34. See, e.g., Michael Boylan, The Weapons Continuum, N.Y. TIMES (DEC. 18, 2012), http://opinionator.blogs.nytimes.com/2012/12/18/the-weapons-continuum/ ("One condition for a citizen entering a society is the trust that societal institutions will protect her and run the society fairly. If one is unhappy with the current state of affairs, the legitimate venue is the ballot box and not a rifle or some other weapon."); see Barack H. Obama, President, United States of America, Remarks by the President at Sandy Hook Interfaith Prayer Vigil (Dec. 16, 2012) (transcript available at http://www.whitehouse.gov/the-press-office/2012/12/16/remarks-president-sandy-hook-interfaith-prayer-vigil) ("[T]his job of keeping our children safe, and teaching them well, is something we can only do together, with the help of friends and neighbors, the help of a community, and the help of a nation. And in that way, we come to realize that we bear a responsibility for every child because we’re counting on everybody else to help look after ours["]);
indulgence, on the one hand, and those for whom the use of even more technology to avoid environmental limits on human productivity and growth expresses the signature virtue of our species. Whose facts the law credits thus signifies whose side it is taking in contests over the best life, magnifying the pressure, conscious and unconscious, to defend the position of one’s own group.

This is a recognizable feature of public reactions to decisions of constitutional law as well. Individuals who form strong, group-reinforcing perceptions of the facts and issues in a culturally charged controversy—one involving, say, the conduct of anti-abortion or civil rights protestors alleged to have crossed the line between persuasion and intimidation; or perhaps the behavior of the police in a suit involving alleged police brutality; or maybe a ban on some minority religious practice (ritualistic drug use; polygamy) asserted to be contrary to public health—are shocked when the Supreme Court or some other tribunal disagrees. The evidence is plain to see; only bad faith—or blind partisanship—can account for the court’s profession to see things otherwise. Those whose group allegiances are affirmed (as it were) by the decision respond with outrage toward the other group, whose (plainly) incorrect depiction of what was going on in the case evinces their refusal to abide by constitutional rules of law.

Such conflicts raise the stakes of constitutional adjudication—or rather simply change them. The outcome of a case upholding a judicially crafted “bubble zone” around an abortion-clinic entrance does not signify merely that this group of anti-abortion protestors “screamed in
the face" of clinic patrons; it denigrates the dignity of all abortion opponents by refusing to uphold their right to advocate a contested vision of the best life "in the face" of public disapproval.39 Or at least it does in the eyes of those who plainly saw the protestors as having been engaged only in "peaceful chanting and singing" akin to that of the Civil Rights era freedom riders.40 The determination that no reasonable jury could fail to see the risk of death to which a fleeing motorist exposed the public stigmatizes as unreasoninable citizens who predictably see the denigration of equality in the decision of the police to chase down an African-American motorist for a minor traffic offense.41 Against the background of culturally polarized reactions to such decisions, constitutional cases are no longer seen as adjudicating the facts of particular disputes, but as determining the status of the cultural groups with whom contending characterizations of those facts have become identified.42

This dynamic itself frustrates the principal of liberal neutrality that animates the Constitution's civil liberties provisions. Judicial enforcement of constitutional rights is supposed to certify that law is not being used as an instrument for advancing a partisan cultural orthodoxy. Such an assurance invests the law with legitimacy—a quality that makes it possible for individuals to reconcile assent to legal obligation with their identities as autonomous moral agents and thus, just for the state to demand such assent. The distorting effect of cultural cognition on public perceptions of constitutional decision-making subverts legitimacy: rather than quieting anxiety over state neutrality, the enforcement of constitutional law itself multiplies the occasions in which the adherents to competing moral outlooks experience law as denigrating their cultural identities.

This effect, too, is a testament to a deficit in both constitutional theory and craft. Our system of constitutional enforcement insists on judicial reason-giving precisely because of the contribution it makes to securing legitimacy. It is simply assumed, however, that conveying the neutrality of the law requires doing nothing more than recounting, in the professional idioms familiar to lawyers, how doctrines aimed at

producing neutral results were brought to bear on the facts of the case. I am, as I indicated, unsure whether and to what extent the reasons that judges offer in their opinions reflect the biasing influence of cultural cognition on their own perceptions. But it seems indisputably clear to me that how they are explaining themselves is doing nothing to counteract the predictable, and predictably illiberal, impact of cultural cognition on public perceptions of what courts are doing.43

Going back to They Saw a Game, it is as if with every call by the referee, alternating halves of the stadium were convinced the contest had been fixed. The game would not go well, no matter how fairly and accurately the referee was in fact making her decisions.

We—scholarly theorists, judges, and lawyers—know that publicly conveying neutrality is integral to securing the substance of neutrality. But we do not have an informed understanding of how to do that, because our doctrines and practices have remained unacquainted with the sorts of psychological insights featured in the study of public risk perceptions.

V. COGNITIVE ILLIBERALISM IN RISK REGULATION

This deficit, moreover, is reciprocated by scholars of public risk perceptions. Their inattention to the normative insights of liberal constitutional theory creates a major gap in their assessment of the significance of cultural cognition.

If cognitive illiberalism poses troubling issues for enforcement of constitutional rights, then it poses them (morally, whether or not constitutionally; only a poorly trained lawyer would think the latter is all that matters) for essentially every manner of risk regulation. If it defies liberal neutrality to regulate conduct out of conscious animosity toward the values or identity of a disfavored group, then it is illiberal to do that whether the conduct being regulated is the expression of a noxious idea or the carrying of a concealed handgun. The law transgresses the liberal prohibition on imposition of a political orthodoxy when it enforces "on the whole society" standards grounded in sectarian "conceptions of right and acceptable behavior"44—whether the standards relate to sexual intimacy or the manufacture and consumption of genetically modified foods.

43. See Kahan et al., supra note 29.
It is very much consistent with neutrality, and with the injunction against enforcement of a sectarian orthodoxy, to regulate such behavior in order to advance "legitimate government purposes"—such as the avoidance of non-communicative harms, or the promotion of other goods of value to persons irrespective of their cultural outlooks. But it is not genuinely consistent with these liberal principles to enforce such regulations when perceptions of their contribution to secular ends are unconsciously motivated by resentment of behavior that denigrates sectarian conceptions of virtue. That is cognitive illiberalism—whether the decision makers reasoning in this biased fashion are jurors and judges applying the Constitution, on the one hand, or legislators, administrative regulators, or individual voters enacting laws, on the other.

Unlike constitutional law theorists, many risk-perception scholars are cognizant of how values shape public perceptions of risk and related facts. Moreover, some of these scholars have had important things to say about this phenomenon: for example, that cultural cognition furnishes yet another ground for delegating regulatory power to politically insulated experts who (it is assumed) are impervious to this and every other species of cognitive bias that constrains public rationality; or, alternatively, that cultural cognition is an element of a richer, expressive form of rationality that members of the public use to assess risk and that experts obtusely disregard when they engage in welfarist cost-benefit analysis. Yet none of these positions tell us anything about how to address the problem of cognitive illiberalism.

The latter "expressive rationality" position does not. For one thing, this view elides the question whether the contribution self-defining values make to individuals' risk perceptions are ones that advance their own goals. People have myriad values, the relationship between which are complex and highly context- and role-specific. As a result, the assessments they are likely to make of the impact of cultural cognition are likely to be complex and highly context- and role-specific, too.


46. See, e.g., John O'Neil, Alan Holland, & Andrew Light, Environmental Values (2008); Ackerman & Heinzerling, supra note 5; Dake, supra note 5; Sabine Roeser, Nuclear Energy, Risk, and Emotions, 24 Phil. & Tech. 197 (2011). See generally Kahan & Slovic, supra note 6.

47. See Kahan & Slovic, supra note 6, at 166-68.
It is plausible, for example, to think the students in the *They Saw a Game* experiment would not have been surprised or the least disappointed had experimenters disclosed the results of the study to them. On the contrary, they likely would have viewed the collaboration of their perceptive faculties and their affective group ties as a benefit in that setting. Seeing things from a partisan point of view (up to a point, anyway) enables a form of shared experience and identity that "home town" fans (and others, too, including family members) enjoy.

But imagine one of the students grows up to become a referee and at the end of a particular season is confronted by league officials with evidence that she has been erring in a manner that systematically benefits the team that represents her alma matter. Such news would likely come as a rude and distressing shock. The mechanism would be the same—identity-protective motivated cognition—but its consequences would now likely be ones that this person intensely resents because of the incompatibility of them with her ends as an impartial referee. She will likely want guidance on how to establish an order between her perceptions, on the one hand, and her diverse moral commitments, on the other, that fits her goals.

By the same token, the impact of cultural cognition on their perceptions of risk would likely disturb many lawmakers, regulators, and ordinary citizens. If they are committed to liberal principles of government, then they too will resent the impact it is exerting on their perceptions of the hazards posed (or not posed) by nuclear power and private gun possession, or by recreational drug distribution and public smoking bans. They will see this effect as a bias, not because it is "irrational" to apprehend risk through value-pervaded affective faculties, but because in this setting the values that are informing their perceptions are not the ones that fit their goals and purposes. If they turn to the "expressive rationality" position on risk regulation for help—for guidance in recognizing whether their perceptions are being shaped by this dynamic and for guidance on how to assure a proper ordering between their modes of apprehending the world and their own ends—they will get no response.

Of course, they will not turn to the "expressive rationality" position for this sort of assistance if they do not view the impact of their cultural outlooks on their perceptions of these sorts of risks as inappropriate. They might, in fact, conceive of risk regulation as akin to a game between opposing cultural teams competing to infuse the law with their own partisan vision, and in that case conceive of their own
role as legislators, regulators, or ordinary citizens as akin to that of players or fans. That is, they might not be committed to principles of liberal democratic government at all.

At that point, those of us who are committed to liberalism will be eager to know what sorts of legal procedures and deliberative norms can be used to rebuff those who want to use risk regulation as an instrument for enforcing a moral or political orthodoxy. If we turn to "expressive rationality" for guidance, it will have nothing to say to us.

Nor will we get any help from the view that cultural cognition is "just another bias." The sort of bias cultural cognition is (or at least sometimes can be) is not like the panoply of others catalogued in behavioral economics. "Base rate" and "denominator neglect," "hindsight bias," "representativeness," "status quo bias," "hypermoblic discounting," etc.48—all of these dynamics constrain information processing independently of cultural and political outlooks. Accordingly, whatever distortions they create in public risk perceptions do not divide citizens on cultural grounds.49

This matters because cultural polarization is the medium through which the impulses of cognitive illiberalism are transmitted and amplified. Conspicuous divisions of opinion impart a salience to risk issues that draws attention to and sharpens the antagonistic meanings involved. They also are the source of the reciprocal suspicion of each side that the other is acting in bad faith, and is in fact using the language of risk and related consequentialist idioms as a pretext for enacting laws to affirm its partisan worldview and denigrate that of its adversaries. It is in this climate that factual disputes over whether the earth is becoming warmer, whether more guns mean more crime or less, etc., become "battles for the soul of America"—divisive forms of status competition, the simple existence of which deprive citizens of the assurance they are due that in assenting to legal obligation they are not submitting to a cultural orthodoxy.

The standard "behavioral economic" prescription for dealing with "cognitive biases" in risk regulation—delegation of authority to expert administrators—does not do anything to address this dynamic. Let


us assume (of course, it is only an assumption) that politically insulated regulators will be immune to cultural cognition as they are (supposedly) to other forms of bias. Let us assume, too, that as a result whatever determinations of fact they make will be more responsive to the best available scientific evidence than would be the determinations made by popularly accountable decision-makers on the same issues. There will still be a problem: because dynamics of cultural cognition would continue to influence how citizens perceive such risks, there is no reason to think that expert administrative resolution would be any more successful in quieting cultural polarization over these issues than judicial resolution is in quieting polarization over issues of constitutional law. Indeed, we can anticipate that the accuracy and neutrality of administrative risk regulation, like the accuracy and neutrality of judicial decisions in constitutional law, would themselves simply become another occasion for divisive forms of illiberal status competition.

Indeed, we already see exactly this. The determinations of even "independent agencies" such as the EPA, the NRC, and the FDA, are fonts of culturally divisive risk controversy, as are the decisions of courts that review their regulations.

It is sensible for behavioral economics to try to assimilate cultural cognition. Without it, that theory is at a loss to explain cultural polarization in public risk perception.51

But using cultural cognition to plug this descriptive hole in the behavioral economics account only exposes an equally embarrassing normative one: its failure to address the problem of cognitive illiberalism. Those who study risk regulation need to learn something from the insights and objectives of liberal constitutional theory, too.

VI. DE-BIASING STRATEGIES

I have explained the special sense in which cultural cognition is a bias. The tendency of people—whether as judges or jurors, as legislators or regulators, or simply as democratic citizens—to give evidence the weight that fits their cultural predispositions does not merely skew the accuracy of their perceptions of risk and related facts; it also undermines liberal political principles of neutrality in constitutional adjudication and in risk regulation generally. It is in reference to this effect, then, that "de-biasing" strategies need to be formulated and evaluated.

51. See Kahan et al., supra note 38.
I will now sketch out what some of those strategies might look like. And I really do mean sketch. I admit that I am just not able, based on what I know or what I can reliably extrapolate, to identify measures that I am confident will "solve" the problem I have described. Indeed, the motivation for this article is a compound of the urgency I feel to promote comprehension of the problem of cognitive illiberalism and the perplexity I experience in assessing what to make of it and how to address it. My sketch, then, is intended as part illustration and part provocation: I want to present an account sufficiently suggestive to communicate the properties de-biasing strategies should have and then hope the manifest incompleteness of what I am describing will inspire others to try to do better. Doing better, moreover, means not only coming up with refinements or additions to these strategies—but also crafting and carrying out real empirical tests that can help to figure out what really works.

I will discuss four prescriptions. The first two are in the nature of decision-making procedures intended to counteract cultural cognition's effect on the determination of legally consequential facts in adjudication. The second two are discourse strategies that aim to lessen the tendency of public justifications for legal or regulatory outcomes to stimulate illiberal status conflict.

A. Juror "Self-Affirmation" Instructions

Cultural cognition is essentially a form of "identity self-defense." When people are confronted with evidence or arguments, their assent to which could diminish the standing of their group or their status within it, people adopt a closed-minded, dismissive cognitive posture.

One remedy for this effect generally is personal "self-affirmation." Before being confronted with evidence that challenges beliefs predominant in or associated with their group's position on a contested issue, individuals are induced to reflect on information that reflects positively on them; a common procedure is simply to have individuals list a set of activities they are good at or desirable traits that they possess. This experience generates a boost in self-esteem that buffers the threatening implications of contemplating the possibility that one's position on an issue important to one's identity might in fact be incorrect. As a
result, studies suggest, self-affirmed individuals react to evidence that challenges such a position much more open-mindedly.52

Such a device could, I suspect, be easily adapted to jury deliberations. In cases in which there seems to be a real possibility that cultural cognition will bias juries—ones like the controversy modeled in the They Saw a Protest experiment—jurors could be instructed to engage in the sort of self-affirmation task I just described, either in the course of jury selection (when it is common for members of the venire to supply various forms of written information) or at the outset of deliberations after the presentation of evidence in the case has been completed.53

B. Moderation-Forcing Deliberation Procedures

On culturally divisive issues, it is commonplace for individuals to overestimate how uniformly and strongly held a position is, both among members of their own group and among members of the opposing one. This effect can magnify the divisive effects of cultural cognition in deliberative settings. Worried about the potential reputational impact of equivocating on their own group’s position, individuals who are genuinely ambivalent will hold their tongues, conceding the floor to group members who hold less qualified views. This skew will reinforce the impression of those in the group that opinion is stronger and more uniform than it is within their group, further stifling expressions of uncertainty or dissent. As a result, members of the opposing group will perceive, too, that those on the other side are all of one mind—and are in fact unreasoning and intransigent. This perception will trigger additional pressure within that group, whose ambivalent members will also likely be keeping a low profile, to adopt a harder line—thereby reflecting back to the first group the signal that members of the other are biased. In the resulting climate of suspicion and recrimination, the psychic pressure to read the evidence in a manner that fits one’s own group’s position becomes even more intense.

A potential remedy is “deliberative moderation-forcing.” The idea is to make it a requirement that all the participants speak, and in addition to setting forth their position also identify what they regard as the


53. See Kahan et al., supra note 17.
strongest arguments on the other side. Such a procedure, studies find, reduces polarization by dissipating the self-reinforcing misperceptions that fuel it. Afforded, in effect, a procedural immunity, individuals who are equivocal about their group’s position will feel unafraid to speak up. Others in the group will thus observe that there is in fact more ambivalence in their own ranks than they might have anticipated, and thus feel even less pressure to conceal their own equivocation—or to stifle it through defensive cognition. Members of the opposing group, too, will also be spared misleading evidence of how resolute and how unreceptive the other side is to the alternative point of view. They will likely reciprocate with greater open-mindedness on their part.

This device too could likely be adapted to legal settings. Certainly the instruction that individual jurors each speak and acknowledge the strongest counterargument to their own initial inclinations could be made part of the instructions on how juries should conduct deliberations. Multi-member appellate courts could self-consciously incorporate this procedure into their deliberations as well.

C. Aporetic Judging

Judicial opinions tend to be remarkably not evenhanded. An opinion writer can be counted on to insist that every relevant consideration points unambiguously, indisputably, inexorably toward the same conclusion—even in the face of an equally one-sided dissent or (as is typical for cases in the U.S. Supreme Court) a division of authority among other courts. There is a well-known element of theatre and ritual here. Indeed, the closer and more controversial those with knowledge recognize a question to be before the decision is issued, the more likely courts are to resort to exaggerated certitude in an effort to “sell” their rulings to a more remote, less informed public.

This practice is a testament to how far removed judicial craft norms are from the insights of social psychology. Although intended to preempt second-guessing, the overstatement characteristic of judicial opinions can be expected to breed exactly that when it interacts with cultural cognition. In cases fraught with contested cultural meanings, members of the public are themselves likely to perceive genuinely complicated factual and legal issues as admitting of only a single defen-

sible disposition—namely, the one that's congenial to their worldviews. A contrary ruling will thus indeed arouse their suspicion. But, rather than mollifying them, an opinion that steamrolls over their position—treating their arguments as essentially frivolous—will fill them with outrage. They will see the aggressively overconfident tone of the opinion as intended to deceive, and thus as incontrovertible proof that the court did not engage the issues with the requisite neutrality. Their (predictable) attack of the court will in turn be viewed as proof of their bad faith by those motivated to form the view of the case affirmed by the court.

In ancient Greek philosophy, *aporia* refers to a style of argumentative engagement that evinced comprehension of an issue's ineradicable complexity. Plato's dialogues, for example, use the aporetic device of serial arguments, the interrogation of which reveal not false but necessarily partial accounts of concepts like justice. At the same time, *aporia*, in the classic sense cannot be simply equated with a state of uncertainty or equivocation; one can reach a decisive conclusion about a problem and still be *aporetic* in assessing it. But if one advances a position that denies or purports to dispel the difficulty that a truly difficult issue poses, or that fails to recognize the genuine weight of the opposing considerations on either side, then one is not being *aporetic*. Indeed, in that case, one is actually not getting the issue at hand, no matter how one resolves it. If a problem presents a genuine tension or conflict among important goods, the absence of *aporia* signifies an aversion to confronting complexity that marks a deficiency in intellectual character.55

An *aporetic* style of opinion writing—one that recognized rather than effaced genuine complexity56—might well help to counteract the self-reinforcing dynamics of group polarization associated with cognitive illiberalism. Frank acknowledgments of the force of competing arguments would remove the assaultive elements of overstatement: its implicit belittling of the intelligence or character of those whose position is being rejected, and its aroma of deceit and bad faith. Indeed, like deliberative moderation-forcing, an *aporetic* mode of engagement would reveal to those who are surprised and disappointed by a legal outcome that the decision maker was not blinded by partisanship and


in fact understood and considered their points. Likewise, it would signal to those whose perceptions are affirmed by the outcome that those who feel otherwise do not lack legitimate bases for seeing things their way. Such an atmosphere, one might hope, would help to extinguish the reverberations of reciprocal recrimination before such a cycle gets started.57

These are admittedly conjectures. They merit, and indeed have already started to receive, empirical investigation.

But the contribution that aporetic opinion-writing can make to addressing cognitive illiberalism, I would argue, does not depend entirely on what such investigations conclude about the impact of such a style of reasoning. I do strongly believe that it is possible to reduce the frequency of the occasions in which citizens apprehend resolution of constitutional issues as declaring winners and losers in illiberal forms of status competition between contending cultural styles. But it seems implausible to imagine, given the nature of the issues involved and the precision with which ameliorating strategies can be expected to operate, that constitutional law can be ever be entirely purged of such resonances. In controversies that continue to bear such significance, then, judges will continue to bear an obligation to carry out their responsibilities in a manner that respects citizens' entitlement to avoid being forced to see their assent to legal obligation as equivalent to acquiescence in a state-endorsed moral orthodoxy. An opinion writing style that insists on gratuitously compounding the disappointment of defeat with the insult of stigmatizing the losers' way of seeing the world as devoid of reason does not satisfy this duty.

D. Expressive Over-Determination

My last point is that judges have a duty to talk in a way that contributes to citizens' stake in being able to conceive of their legal obligations as respectful of their cultural outlooks. So does everyone else.

A prominent theme in contemporary political theory treats the practice of liberal self-government as dependent on "discourse norms." These norms place constraints on the kinds of justifications that political actors can advance for law so that citizens are not forced to conceive of their duty to obey as predicated on the state's endorsement of

57. See Kahan, supra note 34.
goods that presuppose only one understanding, or one set of understandings, of the best way to live.58

Rawls's idea of "public reason"—which consists in those forms of moral justification for public policy that can be found within the "overlapping consensus" of "reasonable comprehensive views"—is the most famous liberal discourse-norm strategy.59 But there are myriad variations, which in one way or another adjust and refine Rawls's particular (some might even think peculiar) understanding of who is constrained to resort to public reason when, along with how they must speak in order to satisfy this duty.60

I now want to offer an account in this spirit. The justification for it in relation to Rawls's concept of "public reason" or in relation to other discourse-norm positions is not that it embodies a philosophically superior understanding of liberalism. Rather it is that the account is inspired by a more psychologically realistic comprehension of the occasions in which the experience of liberal neutrality is genuinely threatened by the styles of discourse that political actors use. Like liberal constitutional theory, liberal political theory, too, I believe, is hampered by its general unfamiliarity with psychology. A form of liberal theory that was better informed by psychology might not generate an account exactly like the one I will describe—and in any case, it would generate something more comprehensive, and concrete. But it would generate an account that tries, as mine does, to understand how discourse norms would operate to address cognitive illiberalism.

One might imagine that the well-being of citizens in a liberal republic depends in part on a healthy discourse environment. That environment consists in the stock of ideas, meanings, memes, traditions, and the like, that enable citizens to conceive of their legal obligations in terms congenial to their group identities. The more plentiful those resources are, the less likely culturally diverse groups are to end up locked into an oppositional state with respect to the meaning of some particular risk or policy-relevant fact. The less likely that it is to happen, the more likely they are make use of the best available scientific

58. See id.
information in forming policies, and the less likely they are to experience the adoption of such policies as gestures of contempt or hostility for their identities.

This assessment supports a discourse norm protective of expressive over-determination. A policy position can be said to be expressively over-determined when it admits to a plurality of cultural meanings congenial to a diverse set of cultural styles. When contributing to public deliberations, political actors—democratically elected officials, certainly, but politically insulated regulators, policy commentators, and ordinary citizens, too—should conduct themselves in a way that enables expressive over-determination and does not upset or enfeeble it.

History furnishes examples of states of expressive over-determination being both adroitly created and recklessly destroyed. In the 1980s, supporters of clean air legislation tapped into the consensus-building energy of expressive over-determination by proposing tradable-emission permits, a regulatory device that simultaneously affirmed communitarian sensibilities favoring collective management of collective goods and individualist confidence in the capacity of private orderings to offset the destructive byproducts of private orderings. For decades, support for social welfare policies was underwritten by a similar mélange of meanings: to egalitarians, such programs were correctives for the arbitrary social disparities created by free markets; to hierarchists, they embodied a societal commitment to protecting the traditional family from economic forces that threatened to drive wives into the workplace to compensate for the depressed income of unemployed or under-employed husbands. That expressively over-determined consensus was dissipated in the late 1960s and early 1970s when a meme that depicted welfare as integral to the economic independence/social dependence of single mothers devoured all competing meanings and imbued opposing stances with identity-defining significance for citizens with these cultural styles.

Today there is a conflict within a conflict over expressively over-determined idioms for engaging climate change. The national climate change debate comprises a ganglia of clashing images—of belching smoke stacks and whirling wind turbines; of grotesquely muscular SUVs and forlorn, shivering polar bears; of Al Gore and Sarah Palin—that bind the positions people take to the group loyalties they owe.

61. See Kahan et al., supra note 49, at 1098-99.
But at least a small part of the discussion is occurring (for the time being, at least) in settings sufficiently removed from the national debate, and sufficiently proximate to other meaning-pervaded domains, to evoke a host of different associations. For example, in coastal states (Florida, Virginia, Louisiana and the Carolinas) and in far-western ones (e.g., Arizona, New Mexico, Colorado, and California), citizens are engaging the issue of adaptation from a plurality of local identities, including "property owner," "utility ratepayer," "municipal resident," "farmer," and "business owner" that cut across the more diffuse affinities in play in the national debate. They are speaking in idioms—ones relating to their shared historical experience, for instance, as people either battered by violent storms or baked by arid, scorching heat—the familiarity and logic of which predate climate change. They are getting their information on the need for action not from culturally iconic figureheads, but from neighbors, local officials, and even their local utility companies.63

Many local actors appreciate how much more hospitable this deliberative environment is to open-minded engagement with scientific information. Consistent with the norm of expressive overdetermination, they have made a concerted effort to decouple debate over proposals for constructing storm surge gates, relocating roads and rail lines, and "hardening" utility assets from global climate change, including (essentially symbolic) proposals for local carbon-emission targets.64

Their efforts have been constrained, however, by the persistent hectoring of those opposed to such a decoupling. Fringe, right-wing populist groups—like the "tea party"—have made sporadic appearances at municipal open-meetings and at administrative stakeholder proceedings to "reveal" the conspiratorial link between adaptation proposals and the "climate change hoax." But significant resistance has also come from environmentalist groups, who cringe at the assent of local advocates to removal of references to "climate change" from state


legislation and local resource plans aimed at promoting adaptation measures.65

Al Gore has been an outspoken critic of President Obama for failing to use the "bully pulpit" of the Presidency to focus attention on carbon emissions when justifying policies such as his Administration's proposal for new national fuel economy standards for new automobiles.66 Many environmentalists have likewise historically derided the rhetorical decoupling of local adaptation from a larger program that includes carbon emission limits as a form of appeasement in a "struggle for the American soul."67

My claim is that the path of expressive least resistance being pursued by the local adaptation forces is morally right, and the obstruction of their progress by the national-debate protagonists is morally wrong. If the goal really is to advance the adoption of sensible policies or promote constructive public engagement with valid policy-relevant science, then clearly it is a mistake to insist that the divisive meanings of the national climate change debate be engrafted onto the local adaptation one. Indeed, such insistence has wrecked attempts to replicate in North Carolina a constructive form of political engagement with climate science now unfolding in states like Florida and Virginia (both of which are controlled by Republican Governors and legislatures) without doing anything to promote compensating forms of engagement with any other aspect of the climate change issue anywhere else.

But my point about who is right and who wrong actually has nothing to do with what communication strategy is most likely to generate any particular set of policies. It has to do with the role of discourse norms in securing liberal principles of government. The climate change localists are engaged in a style of advocacy mindful of the problem of cognitive illiberalism in regulatory politics. By trying to foster a state of expressive over-determination, they are respecting the entitlement of all citizens to be able to engage issues in public debate free of the cog-


nively distorting pressure of apprehending the positions on them as akin to cultural loyalty oaths.

The climate change nationalists, in contrast, are behaving in a manner that recklessly feeds cognitive illiberalism. By insisting on a singular, and singularly divisive frame of meaning, they are needlessly transforming a discussion over how to make people (all people) safe and prosperous into a referendum on the worth of competing ways of life.

If the climate change nationalists care about liberal principles of government, they should recognize that that form of advocacy is inappropriate. If they insist on expressively conflating the local and national debates, despite the barrier that doing so puts in the path of those seeking to help diverse citizens avoid the oppressive choice of assenting to evidence-based regulatory policy and being who they are, then those of us who are committed to political liberalism should condemn them as zealots.

CONCLUSION

In this article, I have proposed the pairing of two bodies of work—the study of risk perception and liberal constitutional theory—each of which supplies the other with the piece it lacks for solving a common puzzle. That puzzle is cognitive illiberalism: the threat that unconscious biases pose to citizens' conscious and genuine commitment to refrain from predicking law on a sectarian orthodoxy. From the study of risk perception, constitutional theory can learn to anticipate the vulnerability of liberal doctrinal strategies to cultural cognition, which motivates individuals to impute frustration of secular goods to behavior that transgresses their partisan understandings of the best life. From constitutional theory, the study of risk can in turn be made to see what makes this dynamic such a significant problem: not merely that it injects error into determination of facts but that it infuses risk with contentious cultural meanings, transforming all manner of regulatory decision-making into occasions for illiberal status competition between opposing cultural groups.

Nothing in the account I have spelled out implies that the enforcement of liberal principles of constitutional law is "psychologically impossible" or that risk regulation is an "inherently" or "inescapably" exercise in cultural partisanship. All it shows the need to combine the best normative understandings that scholars (and jurists) have forged in the former domain with the decision science insights that have be-
come central in the latter. Achieving this reconciliation should be recognized as a critical shared objective of those who study and practice law in both of these fields.