Les ateliers de l'éthique
The Ethics Forum

Democratic Decision Making and the Psychology of Risk

Andreas Christiansen and Bjørn Gunnar Hallsson

See table of contents

Explore this journal

Cite this article

Christiansen, A. & Hallsson, B. G. (2017). Democratic Decision Making and the Psychology of Risk. *Les ateliers de l'éthique / The Ethics Forum*, *12*(1), 51–83. https://doi.org/10.7202/1042278ar

Article abstract

In many cases, the public (or large parts of it) want to restrict an activity or technology that they believe to be dangerous, but that scientific experts believe to be safe. There is thus a tension between respecting the preferences of the people and making policy based on our best scientific knowledge. Deciding how to make policy in the light of this tension requires an understanding of why citizens sometimes disagree with the experts on what is risky and what is safe. In this paper, we examine two highly influential theories of how people form beliefs about risks: the theory that risk beliefs are errors caused by bounded rationality and the theory that such beliefs are part and parcel of people's core value systems. We then discuss the implications of the psychological theories for questions regarding liberal-democratic decision making: (1) Should policy be responsive to the preferences of citizens in the domain of risk regulation? (2) What risk-regulation policies are legitimate? (3) How should liberal-democratic deliberation be structured?

51

DEMOCRATIC DECISION MAKING AND THE PSYCHOLOGY OF RISK

ANDREAS CHRISTIANSEN
POSTDOCTORAL FELLOW, DEPARTMENT OF MEDIA COGNITION AND COMMUNICATION,
UNIVERSITY OF COPENHAGEN

BJØRN GUNNAR HALLSSON
PH.D. STUDENT, DEPARTMENT OF MEDIA COGNITION AND COMMUNICATION, UNIVERSITY OF COPENHAGEN

ABSTRACT:

In many cases, the public (or large parts of it) want to restrict an activity or technology that they believe to be dangerous, but that scientific experts believe to be safe. There is thus a tension between respecting the preferences of the people and making policy based on our best scientific knowledge. Deciding how to make policy in the light of this tension requires an understanding of why citizens sometimes disagree with the experts on what is risky and what is safe. In this paper, we examine two highly influential theories of how people form beliefs about risks: the theory that risk beliefs are errors caused by bounded rationality and the theory that such beliefs are part and parcel of people's core value systems. We then discuss the implications of the psychological theories for questions regarding liberal-democratic decision making: (1) Should policy be responsive to the preferences of citizens in the domain of risk regulation? (2) What risk-regulation policies are legitimate? (3) How should liberal-democratic deliberation be structured?

RÉSUMÉ :

Dans de nombreux cas, le public (ou une grande partie de celui-ci) veut restreindre une activité ou une technologie qu'il croit être dangereuse, mais que les experts scientifiques considèrent être sécuritaire. Il y a alors une tension entre le respect des préférences des gens et des politiques fondées sur nos meilleures connaissances scientifiques. Décider comment élaborer une politique à la lumière de cette tension nécessite de comprendre *pourquoi* les citoyens sont parfois en désaccord avec les experts à propos de ce qui est risqué et ce qui est sûr. Dans cet article, nous examinons deux théories très influentes sur la façon dont les gens forment des croyances sur les risques : la théorie selon laquelle les croyances liées au risque sont des erreurs causées par la rationalité limitée et la théorie selon laquelle ces croyances font partie intégrante des systèmes de valeurs fondamentales des personnes. Nous discutons ensuite les implications des théories psychologiques pour les questions touchant la prise de décision libérale-démocratique : (1) Les politiques devraient-elles être sensibles aux préférences des citoyens dans le domaine de la régulation des risques? (2) Quelles politiques de régulation des risques sont légitimes? (3) Comment la délibération libérale-démocratique devrait-elle être structurée?

1. INTRODUCTION

It is a common and immediately plausible thought that, in a liberal-democratic state worthy of the name, the public should play a substantial role in the policy-making process. It is an equally common and plausible thought that, in an enlightened state worthy of the name, policy making should be based on our best understanding of the relevant facts, which in many domains entails that policy making should be based on scientific knowledge. But now a puzzle presents itself: What to do in cases where the public (or large parts of it) want to restrict an activity or technology that they believe to be dangerous, but that scientific experts believe to be safe (or, conversely, where the public is sanguine about an activity or technology that experts believe to be highly risky)? How, if at all, can liberal-democratic and enlightenment values be reconciled? And if they cannot, how should the two conflicting sets of values be balanced?

In order to answer this question well, we need to understand *why* (parts of) the public sometimes disagree with the experts on matters of risk—we need a cognitive and social psychological understanding of public perceptions of risk. And once we have such knowledge, we need to reflect on what implications the psychological facts have for what role the public ought to play in liberal-democratic policy making. These are our two aims in this paper.

In the first part of the paper (§ 2), we will present and critically assess the evidence for two major and influential psychological theories of risk perception. One is the bounded rationality theory, according to which (nonexperts') thinking about risk is dominated by the use of fast heuristics that lead to predictable biases in risk perception. The other is the cultural cognition theory, which says that lay beliefs about many risks are a result of culturally (or ideologically) biased processing of evidence, and hence are strongly correlated with cultural (or ideological) worldviews. We will argue that, although both theories have their merits, cultural cognition seems to be at play in a majority of the cases where questions of risk regulation are salient politically.

In the second part of the paper (§ 3), we will examine the implications of the psychological theories for three influential liberal-democratic ideas: (3.1) that public policy should be *responsive* to the preferences of citizens; (3.2) that liberal-democratic *legitimacy* requires that policies are reasonably acceptable for all those subject to them; and (3.3) that the public should directly participate in policy making through public *deliberation*. We will focus on claims made by proponents of each of the psychological theories discussed concerning such implications. In particular, we will engage the views of Cass R. Sunstein, on the side of the bounded rationality theory (Sunstein, 2002; 2005; 2006), and of Dan M. Kahan, with a number of coauthors, on the side of the cultural cognition theory (Kahan, 2007; Kahan & Slovic, 2006; Kahan, Slovic, Braman & Gastil, 2006).

On Sunstein's view, the fact that public risk perceptions exhibit the biases characteristic of bounded rationality means that they should be disregarded, and that

policy should instead be determined by the experts using cost-benefit analysis. We will argue that, although Sunstein is right to point out that bounded rationality undermines the case for being responsive to public preferences for risk regulation, his alternative has its own problems.

According to Kahan and coauthors, the fact that risk perceptions are expressions of cultural or ideological worldviews means that they should be treated much as values are treated in liberal-democratic theory. We will argue that this is largely false. However, cultural cognition theory does contain important insights into how we can overcome the conflict between respecting people's values and respecting the truth when making policy concerning risk.

2. PSYCHOLOGICAL THEORIES OF RISK PERCEPTION

Risk perception research has made it clear that there are a number of domains where a substantial proportion of the public disagree with experts about risk-relevant facts. Genetically modified (GM) foods and global warming are two illustrative examples: according to a report by Pew (Pew Research Center 2015), 37% of US adults agree that it is safe to eat GM foods, while the corresponding number among AAAS scientists is 88%. 50% of US adults and 87% of AAAS scientists agree that global warming as a result of human activity is occurring, the latter number increasing to 97% among authors of peer-reviewed articles in climate science (Cook et al. 2013).

The psychology of risk perception aims at explaining such deviations by reference to features of human cognition. The field has been strongly influenced by seminal work by Amos Tversky and Daniel Kahneman on cognitive heuristics and their resulting biases on probability assessments and decision making, as well as their work on prospect theory (Tversky & Kahneman 1974; Tversky & Kahneman 1981; Kahneman 2011). A heuristic is a relatively simple cognitive mechanism that delivers a rapid answer to what may be a complex question, saving time and cognitive resources. While often accurate, the outputs of heuristics may systematically fail under some circumstances. It is these failures that are denoted as biases. So 'heuristic' refers to a cognitive mechanism while 'bias' expresses a normative assessment of the output of this mechanism, to the effect that something has gone wrong from the point of view of a certain normative theory of reasoning (usually probability theory or logic).

To provide an illustrative example: one of the most well-studied heuristics that is also highly relevant to risk perception is the availability heuristic. When using the availability heuristic to answer a question about the probability of an event, people rely on the ease with which they can recall or imagine instances of such events (Tversky & Kahneman 1974). While this may usually yield an acceptably accurate estimate, reliance on the availability heuristic leads to systematic biases in the assessment of probability. The probability of highly salient or widely publicized risks, such as tornadoes or homicides, tends to be overestimated, while the probability of less salient risks, such as heart disease or diabetes, tend to be underestimated (Folkes 1988; Lichtenstein et al. 1978).

Another heuristic whose more recent discovery had a profound impact on the psychology of risk perception is the affect heuristic (Slovic et al. 2004; Finucane et al. 2000). The affect heuristic denotes a tendency for people's judgments of risks and benefits to align along uniformly positive or negative affect towards the risk source. If someone believes that a technology or activity is high risk, she is also likely to believe that its benefits will be low, and vice versa, although there is little reason to suspect that such an inverse correlation usually obtains in reality. This goes beyond people starting with a positive or negative feeling toward a risk source and then generating beliefs about risk and benefits on that emotional background: simply providing people who are naïve with respect to some technology with information that it is high (or low) risk (or benefit) will tend by itself to generate affect, and therefore a belief about benefit (risk) that matches the valence of the initial information. So, if I inform you that a technology, which you currently have no opinion of, is highly risky, this alone will tend to cause you to form the belief that the technology carries little benefit, even in the absence of any direct information about its benefit. More generally, the affect heuristic is representative of an increased awareness within cognitive psychology of the important role emotion plays in risk perception (Roeser, 2010; Slovic et al. 2004).

Heuristic or emotional information processing is typically cast within a dual process framework where it is contrasted with more deliberate, analytical reasoning (Evans 2008; Reyna 2004). When someone is thinking about a technology or activity, a heuristic may yield an initial verdict about risk. Depending on motivation and ability, deliberate reasoning may then be used to scrutinize and possibly override this initial verdict with one that is the result of more deliberate processing (Evans & Stanovich 2013). Heuristics that yield strong intuitions or powerful emotional responses are naturally less likely to be overridden.

2.1. Bounded rationality theory

Psychologists are largely in agreement about the above core findings. Nevertheless, there is substantial disagreement about deeper theories of the psychology of risk perception. We first present bounded rationality theory. The term 'bounded rationality' is sometimes used simply to denote that we as humans are subject to limitations in our decision-making apparatus, compared to an ideally rational agent. This is not controversial. What we call *bounded rationality theory* is a more specific series of claims. It holds that our cognitive apparatus aims at providing accurate factual beliefs, but is fallible in achieving this aim because of overreliance on heuristics. When we form a belief about some risk-relevant fact, the function of that belief is to accurately represent some state of affairs to help us make better choices. However, beliefs may fail to fulfil this function because of cognitive limitations. Subjects may lack the time or processing capacity to engage in deliberate reasoning, and therefore rely on heuristics; and since heuristics are vulnerable to biases, our beliefs may be mistaken. These mistakes can be characterized as "blunders" (Sunstein, 2005): they stem from one's acceptance of the output of heuristic processing and failure to engage in

sufficient reasoning. When lay people disagree with experts about risk, the reason, according to bounded rationality theory, is that lay people often blunder.[1] They rely on heuristic processing, with their associated biases, in their assessment of risk, whereas experts tend to rely on deliberate reasoning including the scientific method and cost-benefit analysis.

Bounded rationality theory has a wealth of research to support it. It rests largely on the literature on core heuristics such as availability, the affect heuristic, framing, and anchoring—which is extensive and well replicated (Klein et al. 2014; Shafir & Leboeuf 2002; Kahneman 2013; Tversky & Kahneman 1981). Additionally, there is some support to the claim that many mistaken beliefs and bad decisions stem from heuristic processing and that increased deliberate processing tends to predict more accurate beliefs and better decisions. One line of research to provide this support is based on individual differences in rational thought (Stanovich & West 1998). People who score highly in one type of test of deliberate reasoning tend to score highly in others (Stanovich & West 2014), and often make better decisions. For example, they tend to make choices under uncertainty that are more utility maximizing compared to people who score low (Frederick 2005). Another approach is to experimentally impair deliberate reasoning through time pressure or a concurrent cognitive load task, or conversely to force a time delay or otherwise attempt to promote reasoning. Inhibiting reasoning consistently leads to errors and to more impulsive behaviour and risk aversion, while bolstering reasoning at least sometimes has the opposite effect (Benjamin et al. 2013).

An aspect of bounded rationality theory that will be important going forward is the implication that people would recognize many of their beliefs as erroneous if they were to engage in the deliberation required to correct their blunder. This hypothetical change of belief might then give rise to different assessments of risk, which would, by virtue of their increased accuracy, be better able to further people's own interests. Thus, adherents of bounded rationality theory can provide a justification for a policy that ignores people's actual beliefs by pointing out that, in addition to better serving their interests, the policy also respects the belief that people actually would have if they were to consider the issue more carefully.

Thus, if the bounded rationality explanation is correct, then we should expect that those parts of the population who disagree with expert judgment about risk-relevant facts do so in part because of a lack of cognitive resources. There are certainly cases where this is borne out. For example, people who tend to rely on intuitive processing profess greater belief in the efficacy of truly ineffectual treatments such as homeopathy to cure disease (Lindeman 2011). However, questioning the general truth of this prediction is at the heart of the cultural cognition critique of bounded rationality, to which we turn in the next section.

2.2. Cultural cognition theory

As mentioned, there is very little disagreement that humans do rely on heuristics and display biases in their thinking about risk.[2] However, the notion that mistaken factual beliefs as a rule are due to the operation of heuristics has come under strong empirical attack from cultural cognition theory. Cultural cognition theory has its roots in anthropological work that describes societal conflict over risk as structured along two cultural dimensions (Douglas & Wildavsky 1983). One dimension, individualism-communitarianism, classifies people according to the extent to which they prefer collective solutions to societal problems over individual and market-driven solutions. The other, egalitarianism-hierarchy, describes the extent to which one prefers firmly stratified social orderings in roles and authority. These two dimensions combine into cultural worldviews, which to a large extent determine people's perception of various risk factors depending on their congeniality or lack thereof to the worldview in question. For example, hierarchical individualists will tend to view regulation aimed at industry as questioning the competence of societal elites and the ability of market forces to solve problems, and therefore tend to view the activity of industry as low risk and not requiring such regulation.

This helps explain a feature of risk perception that is hard to make sense of from within a purely bounded-rationality framework: namely, that attitudes toward many risks form coherent clusters that are sharply divided along political and social fault lines. The above-mentioned figure of 50% of US adults affirming the reality of anthropogenic global warming hides a sharp division within the country: the number is only 15% among conservative republicans, but 79% among liberal democrats (Pew Research Center 2016). Likewise, if one denies the reality of global warming, one is also likely to profess the safety of nuclear power and to favour less gun control. One suggestion from bounded rationality theory might be that this shows one part of the population to be generally more disposed to rely on heuristics than the other. But one would then expect that this group would consistently hold beliefs that are contrary to scientific experts, which is not the case (e.g., as regards the safety of nuclear energy, Pew Research Center, 2015).

To the anthropological base, cultural-cognition theory adds work from psychology on confirmation bias, motivated reasoning, and identity-protective cognition, all of which describe how humans may be biased in their search for, and evaluation of, evidence (Nickerson 1998; Kunda 1990; Dawson et al. 2002). Humans tend to seek out and evaluate evidence in ways that are congenial to their believed or desired conclusions. We tend to accept evidence in favour of our favoured belief with little scrutiny. If the output of a heuristic bolsters a favoured position, then we are unlikely to engage deliberate reasoning to check and possibly overwrite this response. On the other hand, evidence against favoured beliefs is heavily scrutinized and subsequently tends to be deemed weak, while heuristic responses that run counter to a favoured belief will tend to activate deliberate reasoning in an attempt to find an alternative response (Taber & Lodge 2006;

Dawson et al. 2002; Kahan et al. 2017). In evidence-search situations, where people are given the choice between viewing evidence that supports or disconfirms their favoured view, subjects tend to select supporting evidence (Jones & Sugden, 2001; Taber & Lodge, 2006).

So, according to cultural cognition theory, cultural worldviews, not costs and benefits, to a large extent determine people's basic attitudes toward various risk sources. These worldviews furnish us with our basic values, which in turn cause us to engage in motivated reasoning in dealing with evidence, with the aim of reaching factual beliefs about these risk sources that protect and bolster the attitude in line with our values.

This suggests a flaw in the bounded-rationality picture. Mechanisms such as motivated reasoning and identity-protective cognition are not heuristics. They are instances of deliberate reasoning, but instances where the aim appears not to be merely a correct appreciation of the facts, but rather to provide support for a particular conclusion. When cultural worldviews are in play during evaluation of evidence regarding a risk source, we are likely to use our reasoning to assess the evidence such that it comes out supporting the position that confirms our worldview. This in turn predicts that widespread increased reliance on reasoning rather than heuristics will not necessarily bring about convergence towards a view closer to the truth. Rather, we should expect those with the greatest propensity and ability to engage in deliberate processing to be *best* at making the evidence yield their favoured conclusion (Kahan 2013).

In an illustrative study (Kahan et al. 2017), participants were asked to assess which of two conclusions the results of a (fictional) study supported. In the control version of the task, the study in question was on the efficacy of an experimental crème for the treatment of skin rash. The study's results were presented as a two-by-two matrix, with one dimension denoting whether study subjects' rash got better or worse, and the other denoting whether the subjects had received the treatment or the placebo. Each cell contained a number indicating how many people experienced a certain combination of these dimensions (e.g., people whose rash got better *and* who had received the treatment). Participants had to detect correlation between the variables in order to correctly solve the task. This was so difficult that less than half of participants provided the correct answer (i.e., the result was lower than chance), and performance increased with numeracy (a measure of deliberate processing ability) regardless of cultural background.

In the experimental version of the task, the study was on the effect of gun-control legislation on crime. Here, the cells corresponded to cities that had either implemented a gun-control law recently or not, and whether crime had increased or decreased (e.g., one cell contained the number of cities that had not implemented gun-control *and* had experienced a decrease in crime). Here, a sharp division along cultural lines was seen. If given a version where the correct answer was that crime had decreased as a result of gun control, then liberal participants were

likely to find the correct response, and this likelihood increased sharply with numeracy scores. However, conservative participants given this version were very unlikely to find the correct response, and increased numeracy had no effect on their likelihood to do so. The converse pattern was found for the version where the correct response was that crime had increased: conservatives were quite good at finding the correct response, and highly numerate conservatives much more so than less numerate ones, and liberals were bad at finding the correct response, with increased numeracy offering a very limited benefit. That is, increased capacity to engage in deliberate reasoning helped attaining true beliefs only when the evidence was supportive of one's worldview. This suggests that simply providing people with evidence or attempting to engage their deliberative faculty rather than heuristics will do little to correct false beliefs, when these false beliefs are congenial to their cultural worldview. It further suggests that, in general, one should not expect increased deliberative ability to lead to convergence on truth, but rather that one should find the greatest amount of cultural divergence among the most reflective, numerate, and educated.

Research from proponents of cultural cognition theory has borne this out. Across a great many culturally contested domains related to risk, such as global warming, gun control, the HPV vaccine, and fracking, cultural polarization is largest among those with the greatest reflective abilities (Kahan et al. 2012; Kahan et al. 2010; Kahan et al. 2013; Kahan 2015). It thus becomes highly problematic to refer to false beliefs that are the result of the mechanisms described by cultural cognition theory as blunders. In many cases, they may be the result of a large amount of deliberate reasoning, rather than an uncorrected heuristic. Likewise, the notion that policy-makers can assume that people's factual beliefs would align with those of scientific experts if only they were to reflect more becomes untenable. What one could expect is rather that increased reliance on deliberate reasoning would lead to attitude polarization: more extreme versions of current beliefs (Lord et al. 1979; Taber & Lodge 2006).

Naturally, far from all domains of risk are culturally contested. For example, there is no cultural conflict over artificial food colourings or sweeteners, cell-phone radiation, the MMR vaccine, or genetically modified foods (in the US, but probably not in Europe), and in such domains one finds the expected pattern predicted by bounded rationality theory: that higher scientific literacy and reflective capacity increases the likelihood of agreeing with scientific experts, across cultural groups (Kahan 2015). Thus, one can view cultural cognition theory as describing an important class of exceptions to the general bounded rationality framework rather than as providing a full alternative.

It is an important and, to a large extent, unanswered question for cultural cognition theory why and how certain risks become culturally contested and whether this can be reversed: the HPV vaccine apparently became culturally salient only following a series of missteps on the part of its manufacturer (Kahan et al. 2010), and even global warming was not a particularly divisive issue in the early 1990s (McCright, Xiao, & Dunlap, 2014).

3. LIBERAL-DEMOCRATIC DECISION MAKING

We said at the outset that determining the appropriate balance between relying on experts and including lay citizens' views required understanding what causes citizens to sometimes disagree with experts about what things are risky and what things are safe. We have now seen that the answer is, it's complicated. With respect to some risks, the beliefs of (many) citizens are influenced by heuristics and, as a result, exhibit biases. In those cases, those who are the least scientifically literate and who rely the most on intuitive judgment tend to disagree most with the experts. However, for a substantial number of risks, lay opinion is divided along cultural lines. In these cases, agreement with experts is not correlated with scientific literacy or deliberate, careful reasoning—rather the opposite is true. Instead, an individual's beliefs about the riskiness of some phenomenon largely depends on whether that phenomenon is good or bad according to her basic cultural worldview—her basic values. Furthermore, cases where risk debates have become culturally charged are overrepresented among the risks that exhibit the conflict between experts and (some) citizens, which is our subject in this paper.

So what conclusion can we draw concerning risk management in a state that aims to respect liberal-democratic values and to be enlightened? As noted in the introduction, in assessing the political implications of risk psychology, we will focus on claims that proponents of the two theories we have presented have themselves made. We will structure our discussion according to three core ideas in liberal-democratic political theory. First, there is the idea that public policy should be *responsive* to the preferences of citizens—that is, that differences in public opinion should register as differences in the policies implemented. Second, there is the idea that policies should be such that they could enjoy the assent of all those subject to them. This is most famously engendered in liberal and 'public reason' accounts of political *legitimacy*. And third, there is the idea that the public should directly participate through some form of society-wide *deliberation* on policy issues. We will discuss the implications of the psychological theories for each of these ideas in turn. Before doing so, let us state a couple of clarifications and assumptions.

First, when we are talking about people's risk perceptions in a policy-making context, we are not typically talking about pure factual beliefs. Rather, we are typically talking about one of two things: (i) unprompted exclamations (letters to the editor, demonstrations, etc.) to the effect that a certain risk is *serious*, an activity is *dangerous*, or that *something must be done* about a risk, or (ii) support, in one form or another, for proposals to regulate the relevant risky activity (e.g., by expressing such support in surveys, by voting for such policies directly in referenda, or by basing one's vote for representative bodies on the risk-regulation platform of the relevant party or candidate). These are (more or less specific) opinions concerning *what policies should be enacted*—they are *policy preferences*.

Second, we will assume that there is in fact consensus among scientific experts concerning a given risk. Note here that *experts'* views of risk are typically not risk perceptions in the sense defined above (i.e., policy preferences). Rather, they are estimates of the probabilities of various (primarily negative) effects of a policy, such as deaths, other health effects, or environmental degradation. We will also assume that (parts of) the public express policy preferences that are at odds with this consensus, in the sense that the following three propositions are true: (a) the public want a technology or another potentially risky thing restricted, (b) this policy preference is based on a belief that the thing in question is risky, and (c) expert consensus is that the thing is not very risky.

3.1. Responsiveness

While it is fairly uncontroversial that it is an ideal of democratic systems that policies are responsive to the preferences of citizens, it is not clear what this ideal entails more precisely. In particular, it is not clear what 'public preferences' means—it might be public opinion as expressed in polls, the preferences expressed by those citizens who actively engage in political debate, or perhaps the preferences policy-makers perceive to be prevalent in the population (see Manza & Cook, 2002, pp. 631-632). Furthermore, it is not obvious what is required for policies to be responsive to such preferences. Typical explications merely hint at an answer, such as that politicians should *take preferences into account* or that policy should be *influenced* by public preferences (Brooks & Manza, 2006, pp. 474-475). How preferences should be taken into account or how much they should influence policy is left open—although most agree that "a perfect correspondence" is neither required nor desirable (Gilens, 2005, p. 778). We want here to set aside debates about what responsiveness is or should be. Instead, we focus on a more basic issue—namely, whether there is even a *prima facie* requirement that the policies of a democratic state should be respon-sive to citizens' risk perceptions when these are in apparent conflict with expert beliefs.

3.1.1. Sunstein

Sunstein can be seen as arguing that there is no such *prima facie* requirement. At least, he argues that citizens' policy preferences with respect to the regulation of risk-creating activities should play a relatively limited role in policy making. As an alternative, he argues that a major role should be given to cost-benefit analyses performed by experts in regulatory agencies. More precisely, he supports the current (as of 2016) United States system, in which a central agency of the federal government (OIRA, the Office for Informational and Regulatory Affairs) has a mandate to review and reject, on the basis of cost-benefit analy-ses, regulations suggested by the various technical agencies dealing with envi-ronmental, health, and safety policies (such as the Environmental Protection Agency or the Occupational Safety and Health Administration). A main reason for this is a belief that the technical agencies' regulatory priorities reflect public risk perceptions, rather than scientific estimates (Sunstein, 2002, p. 53, citing

Roberts, 1990). The details of Sunstein's proposals are complex, but the main underlying idea is that policy need not be responsive to public risk perceptions, since on his view these are largely (as we have seen above) the products of cognitive biases of various kinds. This conclusion he derives from a general principle: "democratic governments should respond to people's values, not to their blunders" (Sunstein, 2005, p. 126). Since risk perceptions are based on blunders, democratic governments are not required to be responsive to them.

Is he right about this? One possible reason to think that he is not arises if one thinks that the general principle—that democracies should respond only to values, not to blunders—is false. But it is an open question what it would mean for the principle to be false, since it is unclear what the principle says. The problem is that "values" and "blunders" are not exhaustive of the possible descriptions we may give of people's psychological attitudes. True factual beliefs, for example, are clearly neither values nor blunders. Sunstein's principle, then, says that policies *should* be responsive to people's normative beliefs, but *need not* be responsive to their false (or perhaps only obviously false) factual beliefs. This leaves entirely open what we should do when different people or groups hold divergent factual beliefs, none of which is clearly false. In other words, Sunstein's principle has nothing to say about the criteria for selecting which factual beliefs, beyond the clearly false ones, should be allowed to play a role in policy making.

A natural solution to this problem is to add in a principle for selecting respectable factual beliefs. One plausible such principle, congruent with the ideal of enlightened decision making we mentioned in the introduction, would be to use science as a standard-setter. On such a view, any belief conflicting with the scientifically established facts is not entitled to democratic responsiveness. There are ways of questioning this principle, and especially ways of questioning whether (and how) it could be justified given standard understandings of public reason and the nature of factual disagreements (see, e.g., Jønch-Clausen & Kappel, 2015; 2016). However, we believe the price of giving it up is exceedingly large; since the scientific method is the best known way of generating true factual beliefs, it seems that denying that science can act as gatekeeper for beliefs is tantamount to giving up on having any standards of right and wrong in the empirical domain. So we will accept that beliefs in conflict with established scientific fact are such that democratic governments need not respond to them.

An important caveat needs to be added. In a number of cases, among which are many that are policy relevant, scientific knowledge comes with sizeable uncertainties attached. This needs to be taken seriously by policy-makers. Uncertainty, in effect, means that a number of states of affairs are consistent with the available evidence. In the case of risk, a plausible (but perhaps too simple) way of fleshing this out is to assign only an *interval* of probabilities to a given event, rather than a precise probability (for instance, the probability per year of dying from exposure to pesticides may fall in the interval from one in one million to one in two million). In the case of discrete possibilities—for example, whether

gun control works to lower the number of gun-related deaths per year or not—uncertainty means that we cannot believe either discrete possibility very strongly (i.e., the maximum permissible credence for the proposition "gun control works" is relatively close to 0.5). Where uncertainty is involved, the scientific evidence thus does not permit us to give a unique answer to the policy-relevant question—e.g., what the probability per year of dying from pesticide exposure is, or whether gun control works to lower gun-related deaths. Instead, a number of unique answers are possible. It does not fall within the remit of scientific experts to select which of the set of scientifically permissible unique answers to use.

In many cases, however, policy choice depends on what unique answer is correct in the following sense: if p_1 is true, policy R_1 is required (or preferable), but if p_2 is true, R_2 is required. For example, if gun control works, then gun control is (arguably) required—but if gun control does not work, gun control is not required. In such cases, there is a gap between accepting Sunstein's values-not-blunders principle, and delegating decision-making authority to scientific experts, even granting that 'blunders' includes every belief that is contrary to what science says. Public risk perceptions may play some role in filling that gap.

A more important problem with the values-not-blunders principle is that the risk perceptions of ordinary people, being policy preferences, do not straightforwardly fall on either side of the normative-factual belief divide. Consider how an ideally rational person, of the kind one can meet in decision-theory textbooks, would form her policy preferences concerning a risky activity. Such a person would assign a probability and a value measure ("utility") to each possible outcome of each possible policy, multiply each probability by its utility and sum these products, and advocate the policy that has the highest expected utility. So, even for such a person, a call for a given policy is a consequence of a combination of factual and normative beliefs. Indeed, a policy preference can be made consistent with *any* factual belief, given that the appropriate adjustments are made to the person's normative beliefs. The mere fact that the person calls for a given policy does thus not in itself provide evidence that she has a factual belief that is in conflict with the scientific facts.

However, as we have seen above, the bounded rationality theory that Sunstein relies on provides positive reasons to think that people's factual beliefs concerning risk are often wrong. And (at least to a large extent) the basic fact that nonexperts' beliefs about the magnitude of risks often diverge from the best scientific estimates is not in dispute within psychology. So let us suppose that we can be fairly certain that at least some people have erroneous factual beliefs about the magnitudes of various risks. If it were possible to "implant" true beliefs into such people, then it seems plausible that their risk perceptions (i.e., their more or less precise beliefs about what policies should be enacted) would change.

A very plausible explication of the values-not-blunders principle is then this: what democracy requires is responsiveness to the preferences that people would have had if their factual beliefs were true (or at least not contrary to scientifically

established facts).[3] Call this their *counterfactual fact-based preferences*. In so far as policy preferences that ordinary people express currently—call this their *actual preferences*[4]—are different from their counterfactual fact-based preferences, actual preferences are not the kind of thing democracies need to be responsive to.

The normative appeal of this ideal of policy-responsiveness seems to us considerable (although one might want to consider some minimal criteria for what *normative* beliefs are above board as well). Its main problem is its hypothetical nature. We agree that the ideal form of democratic responsiveness is to the counterfactual fact-based policy preferences of citizens. But in order to implement responsiveness to counterfactual fact-based preferences, we must know (or have reasonably justified beliefs about) what specific preferences a citizen or group of citizens would have had, if they had believed the facts. Note that this is quite a lot harder than having a justified belief that citizens would not have had their actual preferences if they had believed the facts. The real challenge for those who wish to implement responsiveness to counterfactual fact-based preferences is to devise or point to some method for generating reasonably justified beliefs about the specific preferences citizens would have had if they had believed the facts. The only fail-safe way would be to make sure all citizens sincerely believe the facts, to have them determine their policy preferences given those beliefs, and then to make policy responsive to those preferences. But it is of course not possible to run a counterfactual fact-based version of the entire democratic process. So it seems that the best we can aim for is a method that we have reason to believe generates preferences that are reasonable approximations to people's counterfactual preferences.

At least in some places, it seems that Sunstein believes that cost-benefit analysis is a procedure that realizes this. Cost-benefit analysis builds on the approach assumed in decision theory, where (as mentioned above) preferences are a function of separate factual beliefs and value judgments. With respect to factual beliefs, cost-benefit analysis uses the best scientific estimates of the magnitude of risks. As such, it clearly meets the criterion of nonresponsiveness to blunders (although doubts can be had as to whether cost-benefit analysts neglect scientific uncertainty (McGarity, 2002)). With respect to the value judgments, cost-benefit analysts assign a monetary value to a given risk (e.g., a one-in-one-hundred-thousand risk of death per year) based on studies of what people are willing to pay to avoid such a risk, or of what they demand to be paid in order to accept bearing such a risk. Typical ways of measuring willingness-to-pay are studies of wage differentials between risky and safe jobs, and surveys asking people directly for their valuations. Sunstein suggests that "the governing theory" behind this approach "follows [people's] own judgments about risk protection" (Sunstein, 2014, p. 86). Although he also stresses that the current practice does not fully realize the governing theory—in particular, it does not sufficiently take into account differences in risk valuations across individuals—he seems to believe that the general willingness-to-pay approach measures people's own valuations of a given risk (as he says, "the limitations [of current theory] are

practical ones" (Sunstein, 2014, p. 136)). By combining these valuations with the facts and assuming the framework of decision theory, cost-benefit analysis arrives at the preferences people would have had if they had believed the facts.

The idea that the methods of cost-benefit analysis tracks people's own valuations—their counterfactual fact-based preferences—is not universally accepted. It relies on extrapolation of behaviours in one context, in particular the labour market, to all other contexts, and on assumptions from economics and rational choice theory that are in many ways questionable (see, e.g., Anderson, 1993, ch. 9; Hausman, McPherson & Satz, 2017, ch. 9). Furthermore, the very same biases and heuristics that Sunstein is eager to expel from risk management through the use of scientific estimates are likely to influence people's valuations of risks in willingness-to-pay studies. Finally, survey studies frequently register a large number of so-called *protest valuations* (where people state a willingness to pay either nothing or an implausibly large amount, or perhaps decline to state a number at all), indicating a rejection of the very idea of using willingness to pay as a valuation measure for public goods (Kahneman, Ritov, Jacowitz & Grant, 1993). Such responses are typically disregarded, which suggests that cost-benefit analysis is ill equipped to deal with preferences that are not of the type typically relevant in markets. Thus it does not succeed in capturing the counterfactual fact-based preferences of those who reject treating a given policy domain as appropriately governed by the ideals of a market economy.

The conclusions that can be drawn from the above are limited. We have merely suggested that Sunstein's proposal of delegating much of the policy-making power to scientific experts doing cost-benefit analyses is not plausibly an ideal solution to risk regulation. So even if Sunstein is right that risk perceptions—of the unfiltered kind that are expressed in the various more or less precise calls for risk-regulating policies—are too tainted by their partial source in cognitive biases to be taken into account in policy making, his alternative may not be much better. At least, his alternative does not embody ideal responsiveness (i.e., responsiveness to counterfactual fact-based preferences). It is doubtful that ideal responsiveness *can* be fully realized in practice. It may be the case that the available realizable alternatives leave us with a dilemma: if we make policy responsive to expressed risk perceptions, we will be *overresponsive* to false or unscientific beliefs; but if we make policy unresponsive to these risk perceptions, we will be *underresponsive* to values. In other words, the seemingly simple ideal of responsiveness to values and nonresponsiveness to blunders may be an unattainable ideal. Call this the *responsiveness dilemma.*

3.1.2. Cultural cognition

Kahan and his coauthors argue that cultural cognition theory further undermines Sunstein's approach. Recall first what the cultural cognition theory says about how people form risk perceptions. On the cultural cognition model, risk perceptions are not formed in the way assumed by decision theorists (and by Sunstein)—that is, by combining pure factual beliefs about the numerical magni-

tude of risks (expected deaths, probabilities of ecosystem damage, and the like) with pure normative beliefs about how bad the various possible bad effects of a policy are. Instead, people assess (probably mostly unconsciously) the relationship between a possibly risky activity and their cultural worldview—and thus assess at the same time whether *restricting* the activity is justifiable, or perhaps required, according to their view of the ideal society. Thus, as we mentioned above, hierarchical individualists balk at regulation of industry because it questions the competence of elites (hierarchy) and assumes the inadequacy of market solutions (individualism). Conversely, egalitarians dislike the activity of capitalist industry generally, and thus welcome restrictions. Based on such general assessments of the value of activities and of restrictions on them, people form factual as well as normative beliefs about the risks and benefits of the activity, in a kind of post-rationalization procedure, in which motivated assessment of evidence concerning the effects of the activity and policy is central.[5] Consequently, "citizens invariably conclude that activities that affirm their preferred way of life are both beneficial and safe, and those that denigrate it are both worthless and dangerous," and even the factual aspect of risk perceptions (could they be isolated) "express [citizens'] worldviews" (Kahan et al. 2006, p. 1105).

Kahan et al. argue that cultural cognition theory undermines Sunstein's view in two related ways. First, they claim that Sunstein's strategy of using cost-benefit analysis to realize the values-not-blunders ideal "borders on incoherence" (Kahan et al., 2006, p. 1105). In other words, the fact that risk perceptions are due to cultural cognition means that the cost-benefit approach does not realize the ideal embodied in the values-not-blunders principle. On one reading, this would merely be the claim we have just made: that cost-benefit analysis fails to respect values. But of course this would be completely independent of the cultural cognition theory. The values we have argued are overridden in cost-benefit analysis are ordinary normative beliefs (about the value of a human life, say), not culturally influenced factual beliefs (about how many lives a certain activity will claim). Second, they suggest that "bringing the role of cultural cognition into view severely undermines the foundation for Sunstein's refusal to afford normative significance to public risk evaluations generally" (Kahan et al., 2006, p. 1004). That is, they suggest that acknowledging the role of cultural cognition undermines the case for nonresponsiveness to citizens' actual policy preferences.

How might the fact that people's risk perceptions are shaped by cultural cognition further undermine the cost-benefit analysts' approach and/or strengthen the case for responsiveness to actual preferences? We suggest that cultural cognition points to two different facts that may be important: (1) that the relationships between values (in the form of cultural worldviews), factual beliefs, and policy preferences are not as Sunstein and others assume, and (2) that risk perceptions are rooted in cultural worldviews, and therefore are expressions of citizens' values.

Let us first consider issue (1). Here, the claim would be that the fact that risk perceptions are due to cultural cognition means that they do not behave in ways that Sunstein and others assume—for example, that changes in factual beliefs do not change preferences in the way assumed—and that this undermines the strategy of cost-benefit analysis further and/or strengthens the case for responsiveness to actual preferences. Such a claim could be made in two ways:

(i) Since both factual beliefs and policy preferences are due to the same underlying cause, we should not expect changes in factual beliefs to change policy preferences. As Kahan et al. put it,

> risk perceptions originating in cultural evaluation are not ones individuals are likely to disown once their errors are revealed to them. Even if individuals could be made to see that their cultural commitments had biased their review of factual information … they would largely view those same commitments as justifying their policy preferences regardless of the facts. (Kahan et al. 2006, p. 1105)

On this reading, an individual's counterfactual fact-based preferences are likely to be the same as his actual preferences (i.e., the preference he would hold if he believed the facts is likely to be the same as the preference he currently holds). If that is the case, people's actual preferences are at least a good approximation of their counterfactual fact-based preferences. Thus we have a solution to the problem of how to achieve responsiveness to counterfactual fact-based preferences—namely, to use actual preferences. Or, to put the matter differently, it is not true that responsiveness to actual preferences is overresponsiveness to faulty factual beliefs, since actual preferences are not influenced by factual beliefs at all—faulty or not. Reading (i) would, then, give reason to be responsive to citizens' actual preferences.

Reading (i) faces two problems. The first problem is that the claim that changes in factual beliefs do not change policy preferences seems too strong, and it goes beyond what can be justified by the evidence that the cultural cognition theory relies on. Cultural cognition is primarily a thesis about how cultural commitments lead to biased assessment of evidence, such that one believes the evidence supports the factual beliefs that fits one's cultural commitments best. But it is possible to debias people at least to some degree, and to bring them towards mutual agreement on the facts. And furthermore, there is evidence that such debiasing alters people's policy preferences, bringing previously opposed parties closer together (Cohen et al., 2007). So it seems to us that the fact of cultural cognition does not justify ignoring the problem of overresponsiveness to false beliefs.

The second problem is that, at least in many policy domains, preferences may lose some of their claim to democratic responsiveness if they turn out to be too resistant to the facts. Resistance to changes in factual beliefs may reveal policy preferences to be based in kinds of value judgments that are unacceptable from

a liberal-democratic point of view—e.g., a desire to regulate purely private behaviour (such as sexual behaviour or harmless commercial activities) or worldviews that deny the fundamental equality of all citizens (such as racist or sexist views). If it were the case that citizens' policy preferences would not change regardless of what the facts are, we would at least need to examine the substantive content of those preferences in more detail—and to reserve judgment as to whether those preferences merit democratic responsiveness until we have a better understanding of what that substantive content is.

(ii) Since policy preferences and factual beliefs are both caused by people's cultural worldviews (i.e., their most basic values), any change in factual beliefs requires a change in basic values.

Suppose a given citizen actually has faulty factual beliefs, and that these beliefs are due to cultural cognition. According to reading (ii), the basic values this citizen actually holds are *not* the basic values she would hold in the counterfactual case where she came to believe the facts. The cost-benefit analysts' method is essentially an attempt to disentangle actual factual beliefs from actual value judgments. The analysis then recombines actual value judgments with the *true* facts, and thereby generates a policy preference. But on reading (ii), such an approach does not succeed in revealing citizens' counterfactual fact-based preferences. The cost-benefit method uses a citizen's *actual* values, but cultural cognition shows that these are likely to be different from her *counterfactual fact-based* values. In other words, a citizen's counterfactual fact-based preferences are not (as Sunstein believes) a function of her actual values and the facts, but a function of a *new* set of values and the facts.

Reading (ii) would show that the cost-benefit analysts' method does not successfully track people's counterfactual fact-based preferences. It also suggests that it is difficult to predict how people's preferences would change if they sincerely came to believe facts that are in conflict with their cultural worldviews. Thus it lends support to the use of more deliberative methods, wherein real flesh-and-blood people are allowed to undergo a change in their views in response to facts and arguments (unlike methods like cost-benefit analysis, which seeks to *infer* what people would prefer from data about what they actually believe, value, and prefer). Consequently, the "deliberative debiasing" methods Kahan et al. argue in favour of using are supported by this reading (Kahan et al., 2006, pp. 1100-1104).

Kahan et al.'s other claim—that cultural cognition supports responsiveness to actual preferences—is not supported by reading (ii). At best, reading (ii) shows cost-benefit analysis to be a worse approximation of the ideal of responsiveness to counterfactual fact-based preferences than we might otherwise have thought. However, this merely makes the responsiveness dilemma worse, by making one of the horns of that dilemma worse. It is not obvious, however, that reading (ii) is of much help in deciding how to choose when faced with a responsiveness dilemma—that is, if we have to choose between responsiveness to actually expressed preferences and (something like) cost-benefit analysis.

Let us now move to issue (2), the fact that cultural cognition theory shows risk perceptions to be expressions of values. Kahan et al. state that "when expert regulators reject as irrational public assessments of the risks associated with putatively dangerous activities … they are in fact overriding *values*" (Kahan et al. 2006, p. 1105). It is, unfortunately, not clear what is meant by "public assessments of … risks" in this quote. On the one hand, the phrase might refer to policy preferences, such as that a given activity A is dangerous and should be regulated. On the other hand, it might refer to people's purely factual beliefs about the magnitude of risks. Let us now consider each of these two readings of issue (2) in turn (we call them readings (iii) and (iv) to avoid confusion with (i) and (ii) above):

(iii) *Experts are overriding G_1's values because they implement a policy R_2 that is different from G_1's preferred policy R_1.*

Recall that the kind of case we are interested in has the following structure: (a) the public wants a technology or another potentially risky thing restricted, (b) this policy preference is based on a belief that the thing in question is risky, and (c) expert consensus is that the thing is not very risky. In the group-based framework of cultural cognition, 'the public' should be replaced with some cultural group. So we assume that a cultural group G_1 wants the activity A restricted through policy R_1, and that G_1 wants this because they believe p, that A carries certain risks. The experts, based on sound science, believe $Øp$ (i.e., that A does not carry those risks) and therefore implement a policy R_2 that does not restrict A appreciably.

In cases of this kind, it is hard to see why we should accept that implementing a policy other than R_1 overrides G_1's values. By assumption, G_1 prefers R_2 *because* they believe p—the implication being that they would not have preferred R_1 if they had believed $Øp$ (i.e., that R_1 is not their counterfactual fact-based policy preference). Once more, there are now two possibilities for what G_1's policy preference would then have been if they had believed $Øp$. First, G_1 might have preferred, or at least acquiesced to, R_2, the policy implemented by the experts. In that case, the expert decision procedure would have achieved its ideal aim. Thus there would be no reason to be responsive to G_1's actual preference, and we would have no reason to object to the experts' decision procedure either. Second, G_1 might have preferred some third possible policy R_3. In that case, we would still have no reason to demand that policy be responsive to G_1's actual preferences. However, there would be reason to complain that the experts' decision procedure has failed to be responsive to G_1's values. Insofar as we cannot tell *a priori* whether G_1 would have preferred (or acquiesced to) R_2 or not, the conclusion that follows is that we cannot be confident that the experts' decision procedure is responsive to G_1's values, in the absence of some effort to determine what G_1's counterfactual fact-based preferences are.

But perhaps the assumption that G_1 prefers R_1 *because* they believe p is not correct. That is, perhaps the case is one in which G_1 would prefer R_1 regardless

of the facts—G_1's factual belief that A is dangerous is merely a *post hoc* rationalization of the group's policy preference, which it holds for other reasons than that A is dangerous. Kahan and Braman (2008, pp. 51-54) suggest that it is only in cases of this kind—where people would not alter their policy preference even if they came to believe the facts—that there is a demand for policy responsiveness to preferences. At the same time, however, they speculate that people would *not* be inclined to hold on to their preferences if they were to realize that their factual beliefs are the product of cultural cognition, at least in the case they are discussing (cases of self-defence). The same might well be the case in typical instances of risk regulation. In the case where people *would* hold on to their policy preferences after coming to believe the facts, the problem we mentioned under reading (i) above recurs: G_1's preference for R_1 has some basis other than that A is in fact risky, and that basis may show the preference to be less reasonable than it initially seemed.

Consider, for example, the case of regulation of industry pollution. Recall that hierarchical individualists tend to be sceptical of such regulation because it casts doubt on the competence of societal elites and the ability of market forces to solve problems, and consequently tend to believe that the risks associated with industry pollution are low. But suppose hierarchical individualists were brought to sincerely believe that some industry's emission of a certain chemical C creates severe risks to the health of those exposed, but that they persisted in their policy preference (not to regulate). What could the basis of such that preference then be, other than a blatant disregard for the welfare of those who will likely suffer health problems? A similar problem arises for egalitarians, who are inclined to approve of restrictions of "commerce and industry, which they see as sources of unjust social disparities" (Kahan, 2012, p. 728), and who consequently tend to believe that the risks associated with industry pollution are high. Suppose egalitarians persisted in their desire to regulate emissions of C even after having sincerely accepted that C does not pose a serious risk to anyone. The only possible basis of such a preference is then a general anti-industry agenda. By persisting in their preferences, both the hierarchical individualists and the egalitarians would violate basic norms of risk regulation, such as that people have some right to be protected against serious risks and that harmless private behaviour cannot be restricted.

Thus it seems to us that in the case of risk regulation there is reason to be sceptical of policy preferences that would not change if people were to come to believe the facts. So, while the possibility that policy preferences would not change if people came to believe the facts does provide some reason to be responsive to those preferences, there will simultaneously be a reason not to be responsive. However, in cases where people merely *overestimate* risk (or underestimate, as the case may be), persisting in policy preference is less problematic. It may reflect, for example, a judgment that the aim of protecting people's health is very important relative to the aim of securing favourable conditions for business. But this is just the general problem with cost-benefit analysis we identified above. It is not obvious that the phenomenon of cultural cognition adds much to that problem.

(iv) Experts are overriding G_1's values by denying the pure factual beliefs of G_1 (i.e., p*), since those factual beliefs express values.*

Since believing p is an expression of G_1's values, the validity of G_1's values is denied when expert regulators implement a policy based on the fact that $\emptyset p$ is true. We think the view that merely denying (a group of) citizens' factual views is to be underresponsive to their values has both strange and dangerous implications. Suppose, for example, that the experts in this case implement G_1's preferred policy R_1, but also believe (and state publicly) $\emptyset p$. On the view considered, the implication would be that the experts' policy making is insufficiently responsive to the values of G_1 in this case, even though G_1 got its preferred policy implemented. That seems to us a strange implication, which requires an excessive demand for responsiveness.

Alternatively, consider a case like the one we mentioned above, where G_1 would at least acquiesce to the expert's implementation of R_2 if they were to come to believe the truth (i.e., $\emptyset p$). One might think that, since the belief p is an expression of G_1's values, implementing R_2 exhibits a lack of responsiveness to G_1's values even though R_2 is G_1's counterfactual fact-based preference (or a least would be acceptable to G_1 in those counterfactual circumstances). In effect, this would amount to denying that policy preferences that unequivocally depend on factual beliefs that do not meet the required correctness criterion (i.e., beliefs that are blunders or contrary to scientifically established facts) do not merit democratic responsiveness. This seems to us a dangerous implication. In factual matters, priority must be given to the truth, and to our best methods for finding out the truth. And in fact, Kahan et al. seem to share our worry here. In a response to Sunstein's response to their original paper, Kahan and Slovic "admit to a fair measure of ambivalence about when beliefs formed as a result of cultural cognition merit normative respect within a democratic society," and concede that "if we came off sounding as if we think democracy entails respecting all culturally grounded risk perceptions, no matter how empirically misguided they might be, we overstated our position" (Kahan & Slovic, 2006, pp. 170-171).

In conclusion, Kahan et al.'s scepticism towards Sunstein's proposed use of expert cost-benefit analysis is largely warranted, but it is questionable if the fact of cultural cognition contributes much to the problems with cost-benefit analysis. To be sure, cultural cognition provides a *different* set of reasons for thinking that cost-benefit analysis does not succeed in tracking counterfactual fact-based preferences—but arguably that claim was already very well supported by other reasons. Furthermore, cultural cognition theory provides only very limited reason to be responsive to actual preferences in cases where these are in conflict with experts' scientific assessments of the riskiness of an activity. Cultural cognition theory therefore does not warrant solving the responsiveness dilemma in favour of responsiveness to actual preferences. It does, however, provide support for using deliberative debiasing techniques to solve that dilemma.

3.2. Liberal legitimacy

We now move from the democratic to the liberal aspect of the liberal-democratic ideal—more precisely, to the liberal conception of legitimacy. According to this conception, political power is legitimate only if could be reasonably accepted by all subject to it. While many philosophers are attracted to some version of the liberal legitimacy principle, there is no general agreement on what the principle precisely amounts to. It is controversial how demanding the requirement that political power be acceptable to all is—does it require that all can accept the basic procedure by which laws and policies are made (Rawls's view) or does it require that each law or policy be reasonably acceptable to all? The latter is obviously a much more demanding criterion. It is likewise controversial how demanding the reasonability clause is—should our conception of reasonability be such that the acceptance of most people as they really exist is required, or do we need to secure acceptance only from people whose views meet higher standards of justifiability? And there are more conflicts as well (for an overview, see Quong, 2013).

Kahan et al. suggest that the cultural cognition theory does have important implications for how policy may be made if it is to be legitimate on the liberal conception. On Kahan et al.'s explication of the liberal ideal, it consists in an "injunction that the law steer clear of endorsing a moral or cultural orthodoxy" (Kahan et al. 2006, p. 1106). They then go on to suggest that "it is questionable whether risk regulation should be responsive to public demands for regulation, since these express cultural worldviews"—that is, exactly the kind of views that it would be wrong for policy to endorse according to the liberal ideal. So even though Kahan et al. seem to believe that the dubious factual basis of risk-related policy preferences is not sufficient to strip them of their claim to democratic responsiveness, they suggest that there are *liberal* reasons for making policy nonresponsive to such preferences.

Kahan et al. do not elaborate what they mean by "endors[ing] a moral or cultural orthodoxy." But since they cite the writings of Bruce Ackerman and John Rawls in support of the principle, let us assume that the following, common liberal idea is what Kahan et al. have in mind: legitimacy requires policies to be justified only with reference to reasons that are public, in the sense that all reasonable citizens agree that these reasons count in favour of (or against, as the case may be) policies. Now suppose we have identified an exhaustive set of such reasons, and that these are the only ones actually given weight in the policy-making process. Obviously policies at the same time will reflect factual assumptions about how much various policies realize the values defined by public reasons. If the cultural cognition theory is correct, factual assumptions are not value neutral, since each set of factual assumptions expresses a cultural worldview.

What is the import of this for liberal legitimacy? The basic question is what it means that factual assumptions express worldviews and when that would be a problem. Suppose a policy is justified only on the basis of public reasons and the

facts. In that case, it seems to us strange to say that the policy in question is illiberal merely because the facts are (coincidentally) endorsed by adherents of one cultural worldview. 'Expressing a worldview' must refer to something more substantial than this kind of *correspondence to* a worldview if it is to be a liberal problem. This reflects the basic assumption we endorsed earlier—namely, that the facts, and scientific methods of establishing facts, ought to have priority in policy making.

Perhaps the problem arises only in cases where there is genuine uncertainty about what the facts are. Suppose that the scientific evidence concerning gun control allows for believing either that gun control does prevent deaths from firearm accidents and crimes (call this p) or that gun control does not prevent such deaths ($\emptyset p$).[6] And suppose further that the public reasons bearing on the case are such that if p is true, then gun control should be implemented, and if $\emptyset p$ is true, gun control should not be implemented (e.g., because there is a presumption of liberty). So policy *must* endorse either p or $\emptyset p$, in the sense that one policy follows from p and a different policy follows from $\emptyset p$. Supposing that p reflects the cultural worldview of one group G_1 and that $\emptyset p$ reflects the worldview of G_2, it seems that policy must endorse one group's worldview although the other group's view is not in conflict with science.

Suppose that one thinks that basing policy on either of p or $\emptyset p$ would be illiberal. Such a view would run into the following problem: it is a plausible requirement for any criterion of legitimacy that at least one policy is legitimate. But in the example given here, we must either say that *both* policies are legitimate or that *neither* policy is legitimate, since they are symmetrically situated with respect to their basis in both public reasons and factual assumptions. Since the view that neither policy is legitimate is not a viable option, we must say that both policies are legitimate. Consequently, G_1 does not have a viable complaint that a no-gun-control policy is illegitimate, although it does in one sense express the cultural worldview of G_2—and similarly G_2 has no legitimacy complaint against gun control.

Another possible interpretation of what it means that a policy preference expresses a worldview is that the worldview is the *real, causal explanation* for why a certain person or group has the preference. On this reading, calls for regulation of a given risk, although seemingly justified by reference to public reasons, are really caused by "an unjust desire to use the expressive capital of the law for culturally imperialist ends" (Kahan et al., 2006, p. 1107). Suppose the policy in question is above board in the sense that some combination of public reasons and scientifically acceptable factual assumptions would justify the policy. Would the fact that this legitimate rationale is not the real reason why the policy is implemented constitute a legitimacy problem? The assumption here is that the group implementing the policy sincerely (and correctly) believes that the policy has a legitimate rationale, a fact that they exploit in order to implement a policy that they desire in any case. Such a group could be accused of an unattractive opportunism. But this does not constitute a legitimacy problem on the orthodox inter-

pretations of the liberal legitimacy criterion.[7] The liberal criterion stresses the importance of all groups *being able to* reasonably accept the policy. Since the policy here is *ex hypothesi* justifiable based on a set of normative assumptions and a set of factual assumptions, both of which are reasonable (i.e., the set of public reasons and the set of scientifically accepted facts, respectively), all groups are able to reasonably accept the policy. It would be unreasonable for a group to demand that the factual assumptions best expressing *their* worldview be the basis of the law rather than another set of reasonable factual assumptions.

We conclude, then, that the fact that factual beliefs express cultural worldviews in the way the cultural cognition theory has revealed does not entail any problems from the point of view of the liberal conception of legitimacy in cases where policies are justifiable based on reasonable normative and factual beliefs.

3.3. Deliberation

In the previous section we discussed public reason as (a part of) a substantive account of policies' legitimacy. We were thus concerned with whether a certain class of reasons provide sufficient justification for a policy. But 'public reason' is also frequently used to refer to a certain norm of deliberation. Here, the concern is not so much whether a policy *could* be justified with reference to agreed-upon, public reasons, but what reasons we may make appeal to in the process of policy making—in public and parliamentary debate, in the civil service, and in courts. According to the deliberative norm of public reason, citizens, politicians, judges, and others may appeal only to reasons that are neutral between reasonable conceptions of the good. The idea, then, is to remove all appeals to contested worldviews from the public arena.

Kahan (2007) takes issue with this public-reason norm. On Kahan's reading, the public-reason norm has two rationales: First, it *disciplines* those in power by demanding that they pursue only policies that they sincerely believe are supported by public reasons. And second, it *protects* those out of power by ensuring that laws are such that they can accept them without thereby denouncing their vision of the good life (Kahan, 2007, p. 129). But, according to Kahan, the cultural cognition theory reveals that the public-reason norm fails to produce either of its promised effects. The demand for secular justifications does not prevent those in power from imposing their vision of the good on society, since even the sincerely held belief that a policy promotes the public good reflects a cultural worldview. And the demand does not ensure that political losers accept policies enacted by their opponents either. More likely, they will interpret opponents' arguments for those policies as disingenuous and reflecting a "smug insistence of their adversaries that such policies reflect a neutral and objective commitment to the good of all citizens" (Kahan, 2007, p. 131).

Kahan suggests that the public-reason norm be replaced with its polar opposite, which he calls the "expressive overdetermination" norm. According to this norm, justifications of policies in the public forum should not avoid references to

contested worldviews and conceptions of the good—they should instead attempt to show how the relevant policy promotes the substantial cultural commitments of all groups. Casting this in Rawlsian terms, we might say that the desire for overlapping consensus among adherents of rival comprehensive views should not lead us to ban reference to the content of these comprehensive views—say, to religious values, strongly egalitarian ideals, or free-market principles. Instead we should attempt to show that all of these values, in all their comprehensive thickness, support some policy (Kahan, 2007, pp. 131-132). The proposal builds on research from social psychology on self-affirmation. The kinds of biases in processing of evidence highlighted by cultural cognition theory stem from a motivation to defend one's identity by defending factual beliefs perceived to be important to the groups with which one identifies. Self-affirmation research has shown that these defensive motivations, and therefore the biases, are decreased when aspects of subjects' personal or social identities are affirmed—for example, by allowing them to write a brief essay outlining a value or group membership that is important to them. In effect, affirmation provides an identity buffer such that one can afford to lower one's cognitive defenses. People whose identities have been affirmed are thus more objective in assessing evidence and arguments, either written or during discussions (Sherman & Cohen 2002; Cohen et al. 2000; Correll et al. 2004; Cohen et al. 2007). Expressive overdetermination takes advantage of this: highlighting that a policy is in line with the values of one's group is taken to be one way of affirming the importance of that value. If so, one can expect people to be less biased in assessing the risks and benefits of the policy. Thus, expressive overdetermination is meant to achieve the goals of having public policy recognized by all groups as legitimate, and of diffusing the intensely conflictual nature of politics.

Kahan et al. (2015) provide direct evidence that expressive overdetermination may be effective. Hierarchical individualists were more likely to rate a study concluding that extant emission limits would be insufficient to avoid environmental catastrophe as valid and to express that climate change posed a high risk if they had previously been exposed to a study suggesting that geoengineering was a necessary element in combating climate change. Since geoengineering does not involve imposing restrictions on free enterprise or suggest that corporate elites are unable to solve collective problems, this framing highlighted that the reality of climate change need not threaten hierarchical individualist values. In fact, these values were affirmed insofar as a privately driven use of technology was cast as necessary to combat climate change. This allowed hierarchical individualists to assess the evidence more objectively without threat to their identity.

The realization that seemingly conflict-diffusing mechanisms, such as the public-reason norm, may in fact not work—or may even be counterproductive—seems to us to be the most directly useful insight for political philosophy that follows from the understanding of cultural cognition. Nevertheless, we do have some misgivings about the expressive-overdetermination norm and about Kahan's dismissal of the public-reason norm.

Let us start with the latter. Is it really true that the public-reason norm fails to deliver on both of its promises? First, consider whether the norm disciplines those in power. The cultural cognition theory shows that the mere fact that those in power *sincerely believe* policies to be supported by public reasons does not ensure that policies *are in fact* so supported. However, it remains plausible that the public-reason norm *contributes* to the aim of liberally legitimate policies. The mere demand that evidence that a certain policy promotes publically recognized goods must be produced will likely provide some constraints on what policies will be implemented by conscientious adherents to the public-reason norm. Although processing of evidence is culturally biased as described above, there are limits to the degree to which people can pick the evidence that suits them (Kunda 1990). Furthermore, there is evidence (Vinokur & Burnstein 1978; Luskin et al. 2012; Cohen et al. 2007) that deliberations between adherents of conflicting worldviews or ideologies brings these people closer together with respect to their factual beliefs. Insofar as the willingness (and perhaps even active desire) to engage with the arguments of political opponents is also a part of the public-reason norm, it has resources to diffuse the kind of conflicts that arise from cultural cognition as well.

Second, consider the protective aim of public reason. A corollary of the above is that the public-reason norm does not plausibly increase the likelihood that liberally illegitimate policies will be enacted (rather, it plausibly lowers that likelihood). So there is no reason to think that losers are less well protected under the public-reason norm than in the case where appeals to "thick" values can freely be made. What the cultural cognition theory shows with respect to losers is that they are likely to feel aggrieved even when they have no right to do so (since policies are legitimate). So only if the goal is to ensure *actual* acceptance on the part of losers does the public-reason norm fail. This is a worthy goal, but less important than protecting them from illiberal cultural imperialism.

Now what about the expressive-overdetermination norm as an alternative? Supposing that Kahan accepts the standard public-reason account of *legitimacy*, expressive overdetermination does not contribute to the legitimacy of policies. On that account, a policy that is *in fact* justifiable by reference to public reasons only is legitimate. The fact that a group falsely believes that a policy is not so justifiable does not alter the fact that it is. Furthermore, expressive overdetermination does not contain any resources that increase the likelihood of policies that are in fact legitimate, or any resources that lower the likelihood of policies that are not legitimate.

There are nonstandard accounts of public reason that may be more conducive to seeing expressive overdetermination as having a legitimacy-creating role. On the convergence view of Gerald Gaus, for example, legitimacy requires that each citizen be able to support the policy from within her own total view (Gaus, 2011; Gaus and Vallier, 2009). Gaus's main argument for viewing legitimacy in this way is that reasons that people hold as part of their comprehensive view, but which are not public reasons, may *defeat* the justification of a policy based on

public reasons. Consequently, people would not be able to sincerely accept the imposition of that policy. This line of argument meshes well with the protection function of deliberative norms as Kahan describes it. However, the convergence view faces the potential problem that there will often not be a policy that can gain support from all comprehensive points of view. Additional principles for determining what policies are legitimate in such cases are then needed. Gaus has developed an elaborate theory for this purpose, but nothing Kahan has written suggests that he would go along with Gaus in this regard. If a legitimacy-incurring role for expressive overdetermination is to be grounded in an account like Gaus's, much work remains to be done to flesh out the theory.

Return now to more standard accounts of public reason. Since expressive overdetermination does not contribute to policies' legitimacy, it seems that the expressive-overdetermination norm can be justified only instrumentally, as a means to an end. The most immediately obvious end that the norm serves is to ensure actual acceptance of policies by all groups. And actual acceptance is presumably valuable because it realizes the aims of disciplining the powerful and protecting the powerless. But there is some reason to be sceptical that actual acceptance will realize those goals. Expressive overdetermination can be used to secure acceptance from groups without substantially respecting their values. Consider an example that Kahan points to—namely, French abortion law. This law gives women access to abortions, but in order to secure acceptance from conservatives, this access is available only in an "emergency" (Kahan, 2007, p. 132). However, no criteria for what constitutes an emergency were included, and no questioning of a woman's own declaration that an emergency exists is allowed. In effect, then, the emergency clause is substantively empty, and was included only for its symbolic meaning. While this construction did succeed in creating a consensus on the policy, it is hard to see why those who believe in any serious way that non-emergency abortions is a problem *should* have been satisfied with this law.[8]

On the other hand, expressive overdetermination might be used for another end—namely, to enable people holding conflicting views to converge on the facts (cf. the climate change study described above), and hence to diffuse or avoid cultural conflict over factual questions. Kahan et al. have provided strong evidence that the public-reason norm does not realize this goal particularly well, and that a norm of expressive overdetermination can (perhaps somewhat counterintuitively) realize the goal better. However, and as Kahan himself recognizes, expressive overdetermination is merely one tool for achieving fact convergence.[9]

4. CONCLUSION

We have argued above that the psychological facts of risk perception are complex. Divergences between experts and lay citizens are sometimes at least partly a reflection of lack of scientific literacy and overreliance on heuristics on the part of some citizens. But in other cases, cultural worldviews seem to be

behind differences of opinion over what is risky and what is not. And in fact those seem to be the cases that are most interesting politically, such as global warming, environmental issues, or GM foods (in Europe).

However, we have also argued that the fact that faulty beliefs express people's basic values has few implications for how liberal-democratic states should go about formulating policy with respect to putatively risky activities and technologies. Contrary to what proponents of cultural cognition argue, the fact that risk perceptions express cultural worldviews does not give us stronger reasons than we would otherwise have for making policy responsive to such perceptions. Similarly, the fact that factual beliefs about risks express visions of the ideal society does not undermine the legitimacy of using scientifically accepted facts as the basis for policy making.

This largely means that we are stuck with the responsiveness dilemma that we identified in our discussion of Sunstein's view: if policy is insulated from the people, we risk being underresponsive to citizens' values, and if policy is made in a more populist manner, we risk overresponsiveness to false beliefs. However, the cultural cognition theory does provide some important insights into how this dilemma can be resolved. It supports the case for using structured deliberation methods to determine what citizens' preferences would be if they were to come to accept scientific facts. And it provides significant guidance for those of us who want to reform political discourse in a way that enables reasonable discussion of policies based on common acceptance of the relevant facts.

NOTES

[1] This is a bit of a simplification. Bounded rationality is also consistent with mistakes being due to a lack of information or to social processes such as information cascades or group polarization (Anderson & Holt 1997; Moscovici & Zavalloni 1969; Sunstein 2002).

[2] However, the ecological rationality programme of Gerd Gigerenzer and colleagues points out that, far from being a source of ubiquitous bias, heuristics can often be beneficial, providing "fast and frugal" decision procedures that can rival or even beat analytical approaches (Gigerenzer & Goldstein 1996; Czerlinski et al. 1999).

[3] Discussing the phenomenon of "nudging," where policy proposals have similarly been justified with reference the psychology of heuristics and biases, one observer suggests that the people arguing for such policies "generally believe that social policy should aim to satisfy purified preferences" (Hausman, 2016). "Purified" preferences are preferences people would have had, if they had not been the victims of biases.

[4] Here, and generally in the paper, we use the word 'actually' to indicate what is the case in the *actual* world, as that concept is typically used in possible-worlds ways of speaking of counterfactuals and alethic concepts such as necessity and possibility. That is, we use 'actual' to indicate what is currently the case in the world in which we find ourselves.

[5] There are two likely mechanisms at play: First, people form beliefs about whether a given type of risk regulation is desirable, based directly on their cultural worldview. That is, there is a direct causal link from worldviews to policy preferences. Second, people form *factual* beliefs—about the numerical magnitude of risks—through motivated cognition, wherein cultural worldviews affect people's assessment of the evidence concerning the riskiness (or safety) of an activity. Here, the causal link goes from worldviews to assessment of evidence, and thus to pure factual beliefs, and then in a second step from those factual beliefs to policy preferences. Since pure factual beliefs are not easily disentangled from policy preferences (see, e.g., Kahan & Slovic, 2006, pp. 166-168), it is difficult to test which of these mechanisms is the dominant one. However, in a study of self-defence cases, Kahan and Braman found support for the view that "the influence that values exert over outcome judgments is mediated by the impact of the commitments on individuals' perceptions of the facts" (Kahan and Braman, 2008, p. 45)—i.e., for the second mechanism.

[6] Kahan's own treatment of this case (2007, pp. 120-122) seems to imply that this is the case. However, more recent evidence suggests that gun control does, in fact, lower gun-related injuries and fatalities. See Santaella-Tenorio, Cerdá, Villaveces and Galea, 2016.

[7] There is some debate among theorists of public reason regarding the appropriate role of *sincerity*. Some views within this debate hold it to be a requirement for legitimacy that public reasons are the actual motivation for people's advocacy of a given policy (see Schwartzmann, 2011, pp. 387-390). Kahan et al. may of course defend their position by endorsing such a view, but in doing so they would no longer be able to claim the support of the liberal principle of legitimacy *tout court*.

[8] Of course, one might not think that the anti-abortion party's views were such that they ought to be respected on a liberal view of legitimacy—but the example is illustrative of the risk that expressive overdetermination can be used to manipulate groups to accept policies that illegitimately trample their values.

[9] http://www.culturalcognition.net/blog/2014/4/5/cognitive-illiberalism-expressive-overdetermination-a-fragme.html (comment by "dmk38").

REFERENCES

Anderson, Elizabeth, *Value in Ethics and Economics,* Cambridge, MA, Harvard University Press, 1993.

Anderson, Lisa R. & Holt, Charles A., "Information Cascades in the Laboratory", *American Economic Review*, volume 87, issue 5, 1997, pp. 847-862.

Benjamin, Daniel J., Brown, Sebastian A. & Shapiro, Jesse M., "Who is 'behavioral'? Cognitive ability and anomalous preferences", *Journal of the European Economic Association*, volume 11, issue 6, 2013, pp. 1231–1255.

Blackburn, Simon, *Practical Tortoise Raising: and Other Philosophical Essays,* Oxford, Oxford University Press, 2010.

Brooks, Clem & Manza, Jeff, "Social Policy Responsiveness in the Developed Democracies", *American Sociological Review,* volume 71, issue 3, 2006, pp. 474-494.

Cohen, Geoffrey L., Sherman, David K., Bastardi, Anthony, Hsu, Lillian, McGoey, Michelle & Ross, Lee, "Bridging the Partisan Divide: Self-Affirmation Reduces Ideological Close-Mindedness and Inflexibility in Negotiation", *Journal of Personality and Social Psychology*, volume 93, issue 3, 2007, pp. 415-430.

Cohen, Geoffrey L., Aronson, Joshua and Steele, Claude M. "When Beliefs Yield to Evidence: Reducing Biased Evaluation by Affirming the Self", *Personality and Social Psychology Bulletin*, volume 26, issue 9, 2000, pp. 1151-1164.

Cook, John et al., "Quantifying the consensus on anthropogenic global warming in the scientific literature", *Environmental Research Letters*, volume 8, issue 2, 2013, pp. 1-7.

Correll, Joshua, Spencer, Steven J. and Zanna, Mark P. "An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength", *Journal of Experimental Social Psychology*, volume 40, issue 3, 2004, pp. 350-356.

Czerlinski, Jean, Gigerenzer, Gerd & Goldstein, Daniel G., "How good are simple heuristics?", *in* Todd, Peter M., Gigerenzer, Gerd & ABC Research Group (eds.), *Simple Heuristics that make us smart*, Oxford, Oxford University Press, 1999.

Dawson, Erica, Gilovich, Thomas & Regan, Dennis T., "Motivated Reasoning and Performance on the Wason Selection Task", *Personality and Social Psychology Bulletin*, volume 28, issue 10, 2002, pp. 1379-1387.

Douglas, Mary. & Wildavsky, Aaron B., *Risk and culture: an essay on the selection of technological and environmental dangers*, Berkeley, University of California Press, 1983

Evans, Jonathan St. B.T., "Dual-processing accounts of reasoning, judgment, and social cognition", *Annual Review of Psychology*, volume 59, issue 1, 2008, pp. 255-278.

Evans, Jonathan St. B.T. & Stanovich, Keith E., "Dual-Process Theories of Higher Cognition: Advancing the Debate", *Perspectives on Psychological Science*, volume 8, issue 3, 2013, pp. 223-241.

Finucane, Melissa L., Alhakami, Ali, Slovic, Paul & Johnson, Stephen M., "The affect heuristic in judgments of risks and benefits", *Journal of Behavioral Decision Making*, volume 13, issue 1, 2000, pp. 1-17.

Folkes, Valerie S., "The Availability Heuristic and Perceived Risk", *Journal of Consumer Research*, volume 15, issue 1, 1988, pp. 13-23.

Frederick, Shane, "Cognitive Reflection and Decision Making", *Journal of Economic Perspectives*, volume 19, issue 4, 2005, pp.25-42.

Gaus, Gerald, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World,* Cambridge, Cambridge University Press, 2011.

Gaus, Gerald & Vallier, Kevin, "The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Assymmetry, and Political Institutions", *Philosophy and Social Criticism,* volume 35, issue 1, 2009, pp. 51-76.

Gigerenzer, Gerd & Goldstein, Daniel G., "Reasoning the fast and frugal way: models of bounded rationality", *Psychological review*, volume 103, issue 4, 2005, pp. 650-669.

Gilens, Martin, "Inequality and Democratic Responsiveness", *Public Opinion Quarterly*, volume 69, issue 5, 2005, pp. 778-796.

Jones, Martin & Sugden, Robert, "Positive confirmation bias in the acquisition of information", *Theory and Decision*, volume 50, issue 1, 2001, pp. 59-99.

Jønch-Clausen, Karin & Kappel, Klemens, "Social Epistemic Liberalism and the Problem of Deep Epistemic Disagreements", *Ethical Theory and Moral Practice*, volume 18, issue 2, 2015, pp. 371-387.

—, "Scientific Facts and Methods in Public Reason", *Res Publica*, volume 22, 2016, pp. 117-133

Kahan, Dan M., "The Cognitively Illiberal State", *Stanford Law Review*, volume 60, issue 1, 2007, pp. 115-140.

—, "Cultural Cognition as a Conception of the Cultural Theory of Risk", *in* Roeser, Sabine, Hillerbrand, Rafaela, Sandin, Per & Peterson, Martin (eds.), *Handbook of Risk Research: Epistemology, Decision Theory, Ethics and Social Implications of Risk,* Dordrecht, Springer, 2012.

—, "Ideology, motivated reasoning, and cognitive reflection", *Judgment and Decision Making*, volume 8, issue 4, 2013, pp. 407-424.

—, "Climate-science communication and the measurement problem", *Political Psychology*, volume 36, issue S1, 2015, pp. 1-43.

Kahan, Dan M., Slovic, Paul, Braman, Donald & Gastil, John, "Fear of Democracy: A Cultural Evaluation of Sunstein on Risk", *Harvard Law Review*, volume 119, 2006, pp. 1071-1109.

Kahan, Dan M. & Slovic, Paul, "Cultural Evaluations of Risk: 'Values' or 'Blunders'?", *Harvard Law Review Forum,* volume 119, 2006, pp. 166-172.

Kahan, Dan M. & Braman, Donald, "The Self-Defensive Cognition of Self-Defense", *American Criminal Law Review,* volume 45, issue 1, 2008, pp. 1-65.

Kahan, Dan M. et al., "Who fears the HPV vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition", *Law and human behavior*, volume 34, issue 6, 2010, pp. 501–516.

Kahan, Dan M., Jenkins-Smith, Hank & Braman, Donald, "Cultural Cognition of Scientific Consensus", *Journal of Risk Research,* volume 14, number 2, 2011, pp. 147-174.

Kahan, Dan M. et al., "The polarizing impact of science literacy and numeracy on perceived climate change risks", *Nature Climate Change*, volume 2, issue 10, 2012, pp. 732-735.

Kahan, Dan M, Peters, Ellen, Dawson, Erica C. & Slovic, Paul, "Motivated Numeracy and Enlightened Self-Government", *Behavioral Public Policy*, volume 1, issue 1, 2017, pp. 54-86.

Kahan, Dan M., Silva, Carol L., Tarantola, Tor, Jenkins-Smith, Hank & Braman, Donald, "Geoengineering and Climate Change Polarization: Testing a Two-Channel Model of Science Communication", *Annals of the American Academy of Political and Social Science,* volume 658, issue 92, 2015, pp. 192-222.

Kahneman, Daniel, *Thinking, Fast and Slow*, New York, Farrar, Strauss & Giroux, 2011.

Kahneman, Daniel, Ritov, Ilana, Jacowitz, Karen E. & Grant, Paul, "Stated Willingess to Pay for Public Goods: A Psychological Perspective", *Psychological Science,* volume 4, issue 5, 1993, pp. 310-315.

Klein, Richard A. et al., "Investigating Variation in Replicability", *Social Psychology*, volume 45, issue 3, 2014, pp. 142-152.

Kunda, Ziva, "The case for motivated reasoning", *Psychological Bulletin*, volume 108, issue 3, 1990, pp.480-498.

Lichtenstein, Sarah, Slovic, Paul, Fischhoff, Baruch, Layman, Mark & Combs, Barbara, "Judged Frequency of Lethal Events", *Journal of Experimental Psychology: Human Learning and Memory,* volume 4, number 6, 1978, pp. 551-578.

Lindeman, Marjaana, "Biases in intuitive reasoning and belief in complementary and alternative medicine", *Psychology and Health*, volume 26, issue 3, 2011, pp. 371-382.

Lord, Charles G., Ross, Lee & Lepper, Mark R., "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence", *Journal of Personality and Social Psychology*, volume 37, issue 11, 1979, pp. 2098-2109.

Luskin, Rober C., O'Flynn, Ian, Fishkin, James S, & Russell, David, "Deliberating across Deep Divides", *Political Studies*, volume 62, issue 1, 2012, pp. 116–135.

Manza, Jeff & Cook, Fay Lomax, "A Democratic Polity? Three Views of Policy Responsiveness to Public Opinion in the United States", *American Politics Research*, volume 30, issue 6, 2002, pp. 630-667.

McCright, Aaron M., Xiao, Chenyang & Dunlap, Riley E., "Political polarization on support for government spending on environmental protection in the USA, 1974-2012", *Social Science Research*, volume 48, 2014, pp. 251-260.

McGarity, Thomas O., "Professor Sunstein's Fuzzy Math", *The Georgetown Law Journal,* volume 90, 2002, pp. 2341-2377.

Hausman, Daniel, McPherson, Michael & Satz, Debra, *Economic Analysis, Moral Philosophy, and Public Policy (3rd ed.)*, Cambridge, Cambridge University Press, 2017.

Moscovici, Serge & Zavalloni, Marisa, "The group as a polarizer of attitudes", *Journal of Personality and Social Psychology*, volume 12, issue 2, 1969, pp. 125-135.

Nickerson, Raymond S., "Confirmation bias: A ubiquitous phenomenon in many guises", *Review of general psychology*, volume 2, issue 2, 1998, pp. 175-220.

Pew Research Center, *Public and Scientists' Views on Science and Society*, January 29, 2015.

—, *The Politics of Climate*, October 4, 2016.

Quong, Jonathan, "Public Reason", *in* Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2013 edition)*. URL: https://plato.stanford.edu/archives/sum2013/entries/public-reason/

Rawls, John, *Political Liberalism*, New York, Columbia University Press, 2006.

Reyna, Valerie F., "How people make decisions that involve risk: A dual-process approach", *Current Directions in Psychological Science*, volume 13, issue 2, 2004, pp. 60-66.

Roberts, Leslie, "Counting on Science at EPA", *Science*, volume 249, issue 4969, 1990, pp. 616-618.

Roeser, Sabine (Ed.), *Emotions and Risky Technologies*, Dordrecht, Springer, 2010.

Santaella-Tenorio, Julian, Cerdá, Magdalena, Villaveces, Andrés & Galea, Sandro, "What Do We Know About the Association Between Firearm Legislation and Firearm-Related Injuries?", *Epidemologic Reviews*, volume 38, issue 1, 2016, pp. 140-157.

Shafir, Eldar & Leboeuf, Robyn A., "Rationality", *Annual Review of Psychology*, volume 53, 2002, pp. 491–517.

Sherman, David K. and Cohen, Geoffrey L. "Accepting Threatening Information: Self-Affirmation and the Reduction of Defensive Biases", *Current Directions in Psychological Science*, volume 11, issue 4, 2002, pp. 119-123.

Slovic, Paul, Finucane, Melissa L., Peters, Ellen & MacGregor, Donald G. "Risk as analysis and risk as feelings: Some Thoughs about Affect, Reason, Risk and Rationality", *Risk Analysis*, volume 24, issue 2, 2004, pp. 311-322.

Stanovich, Keith E. & West, Richard F., "Individual differences in rational thought", *Journal of Experimental Psychology: General*, volume 127, issue 2, 1998, pp. 161-188.

—, "The Assessment of Rational Thinking: IQ ≠ RQ", *Teaching of Psychology*, volume 41, issue 3, 2014, pp. 265-271.

Sunstein, Cass R., "The Law of Group Polarization", *Journal of Political Philosophy*, volume 10, issue 2, 2002, pp. 175-195.

—, *Risk and Reason*, Cambridge, Cambridge University Press, 2002.

—, *Laws of Fear: Beyond the Precautionary Principle,* Cambridge, Cambridge University Press, 2005.

—, "Misfearing: A Reply", *Harvard Law Review,* volume 119, no. 4, 2006, pp. 1110-1125.

—, *Valuing Life: Humanizing the Regulatory State,* Chicago, The University of Chicago Press, 2014.

Taber, Charles S. & Lodge, Milton, "Motivated skepticism in the evaluation of political beliefs", *American Journal of Political Science*, volume 50, issue 3, 2006, pp. 755-769.

Todd, Peter M., Gigerenzer, Gerd & the ABC Research Group, *Ecological Rationality: Intelligence in the World*, Oxford, Oxford University Press, 2012.