

Critical Thinking and Cognitive Bias

JEFFREY MAYNES

*Department of Philosophy
St. Lawrence University
23 Romoda Drive
Canton, NY 13617
U.S.A.
jmaynes@stlawu.edu*

Abstract: Teaching critical thinking skill is a central pedagogical aim in many courses. These skills, it is hoped, will be both portable (applicable in a wide range of contexts) and durable (not forgotten quickly). Yet, both of these virtues are challenged by pervasive and potent cognitive biases, such as motivated reasoning, false consensus bias and hindsight bias. In this paper, I argue that a focus on the development of metacognitive skill shows promise as a means to inculcate debiasing habits in students. Such habits will help students become more critical thinkers. I close with suggestions for implementing this strategy.

Résumé: L'enseignement des habiletés associées à la pensée critique est un objectif pédagogique central dans de nombreux cours. On souhaite que ces habiletés soient à la fois transférables—applicables dans un large éventail de contextes—, et durables. Toutefois, développer des habiletés ayant ces deux propriétés est rendu difficile par des biais cognitifs qui sont à la fois puissants et omniprésents, tels le biais de la conclusion agréable, le biais de faux consensus et le biais rétrospectif. Dans cet article, je défendrai l'idée qu'il est prometteur de mettre l'accent sur le développement d'habiletés métacognitives comme moyen d'inculquer des attitudes permettant de combattre les biais. De telles attitudes aideront à développer la pensée critique des étudiants. Je terminerai par des suggestions quant aux manières de mettre en pratique cette stratégie.

Keywords: cognitive bias, critical thinking, metacognition, pedagogy

1. The problem

Developing critical thinking skill is a central educational aim across the curriculum. Critical thinking, it is hoped, perhaps unlike the specifics of course content, is *durable* and *portable*.¹

¹ The portability of critical thinking has been challenged (Willingham, 2007). The metacognitive proposal I offer here shows promise as an answer to this worry about portability based on findings suggesting that metacognitive skill is portable (Schraw, Dunkle, Bendixen, & Roedel, 1995; Scott & Berman,

That is, students' critical thinking skills will last long after they leave the classroom, and will be applicable to a wide-range of subjects. The student who thinks critically in a philosophy class will think critically in a chemistry course, and will think critically in her post-graduate endeavors. As a critical thinker, she will be more successful in her personal and professional life, and be a better citizen.

Siegel (1988) defines a critical thinker as someone who is "appropriately moved by reasons." Ennis (1995) characterizes the critical thinker in terms of a host of dispositions and abilities. Both emphasize that the critical thinker is not only someone who has learned *how* to do something, but is someone who actually *does it*. A critical thinking instructor would not be satisfied to find that her students excelled with critical thinking when in the classroom, but that they left those skills unused beyond its walls.

Teaching critical thinking, however, faces a serious challenge. We are subject to a wide-array of cognitive biases which impede critical thinking, undermining both its durability and portability. A reasoner who is prone to a bias lacks the dispositions to reason critically (at least in those situations where she is prone to the bias). The literature on these biases is extensive (see Kahneman, 2011), and I will only highlight a selection here.

One bias highly relevant to critical thinking courses is *motivated reasoning*, or the tendency to find arguments in favor of conclusions we want to believe to be stronger than arguments for conclusions we do not want to believe (Kunda, 1990). Suppose that I am wealthy, and also a great believer in lower taxes rates for the wealthy. I might argue that such lower tax rates are actually better for others in society through their positive impact on the greater economy. I might even have evidence to support this claim about the economic benefits of such policies. Yet, at the same time, I might find these arguments more compelling than I would if I were in a different position, because they affirm something I already wish to be the case. Motivated reasoning can be subtle, in that the mere presence of arguments or evidence does not suffice to show that we are not affected by it. Instead, it operates on our appraisal of arguments and evidence. It also raises particular concern for the portability and durability of critical thinking skill, as students are more likely to face decisions that involve outcomes they have antecedent convictions about outside of the classroom.

2013; Veenman, Elshout, & Meijer, 1997), though see Kelemen, Frost, & Weaver (2000) for contrary results. With the evidence still coming in on this question, I will not pursue it here.

False consensus bias is the tendency to believe that one's own beliefs are widely shared by others. In one study, Ross et al. asked student volunteers to walk around campus wearing a sandwich board reading "Eat at Joe's" (Ross, Greene, & House, 1977). The participants were then asked how likely their peers would be to also wear the sign. Those who agreed to wear the sign estimated that approximately 65% of their peers would also choose to do so, while those who refused to wear it estimated that only 31% would. That is, the participants expected their peers to make the same decision that they themselves did. This bias interacts with the burden of proof. If I think my own beliefs are obvious and widely shared, I might not see any need to argue for them, or to investigate them closely.

Hindsight bias is the tendency to erroneously see events as inevitable or highly likely after they have occurred. Kahneman and Riepe draw examples from the investment world. "Within an hour of the market closing every day, experts can be heard on the radio explaining with high confidence why the market acted as it did. A listener could well draw the incorrect inference that the behavior of the market is so reasonable that it could have been predicted earlier in the day" (Kahneman & Riepe, 1998, p. 55). A pernicious effect of this bias is that we tend to evaluate judgments and judgment-makers by the outcomes, even if the reasoning was appropriate given the epistemic situation at the time (Kahneman, 2011, pp. 203–204). Suppose a meteorologist has ample grounds to predict a severe storm and does so, but nevertheless, the predicted storm does not come to pass. Under the influence of the bias, I might believe that the nice weather was predictable and so judge that the meteorologist is unreliable. In the future, I might make worse decisions (as I do not trust a reputable judge on weather matters) because I misjudged the epistemic situation of these past decisions.

In addition to these cognitive biases, we are also subject to a metacognitive bias known as the *blind spot bias*, which is the tendency to identify biases more readily in others than in ourselves. We are not only biased in our reasoning, but we fail to recognize those biases. These biases are not attenuated by cognitive ability (K.E. Stanovich & West, 2008), they may even be exacerbated by it (West, Meserve, & Stanovich, 2012). That is, simply acquiring great cognitive ability and sophistication, aims of a general education, will not reduce the effects of the biases, nor will it even make us better aware of how they influence us.

Thagard has argued that critical thinking courses that focus on informal logic do not speak to the actual reasoning processes that students go through. As a result, students are not prepared

to handle these biases and reason according to the normative standards of critical thinking (Thagard, 2011). Worse, it might turn out that these biases are recalcitrant, and pedagogical efforts aimed at developing critical thinking are helpless against them. If so, then we might not only fail to teach critical thinking, but worse, we might simply arm students with sophisticated techniques to rationalize their prejudices and bully opponents.

2. A solution

As instructors, we will not be able to debias our students. What we can do, I will argue, is prepare students to exert some degree of control over their own biases. A student with a well-developed repertoire of strategies on how to control his or her own biases, and with both the knowledge and ability to implement those strategies, will be able to mitigate the effects of those biases. Doing so will contribute to the student's tendency to think critically in the future, addressing a significant challenge to the durability and portability of critical thinking skill.

This approach picks up a challenge recently laid down by Kenyon and Beaulac (2014), who challenge the usefulness of pedagogical strategies that treat debiasing as a self-deployed skill. The pedagogy that they target, which they label the "intuitive approach," is that explicit instruction on the biases will help students to regulate their own biases. Kenyon and Beaulac are right that, given the bias blind spot, we should not expect students to be able to use this knowledge to identify cases of bias and recognize their own biases reliably. In what follows, I offer a strategy to go beyond explicit instruction to help students develop a self-deployed ability to mitigate the effects of their own biases. This ability is responsive to the argumentative context, and not to the self-identification of a bias.

Critical thinking essentially involves *metacognitive* skill, and critical thinking pedagogy should include a focus on developing this skill (see also Maynes, 2013). Typically when teaching critical thinking, we teach cognitive skills, such as argument diagramming or mapping, implicit premise identification, and fallacy identification. The metacognitive skills involved in critical thinking are those skills involved in recognizing *when* these cognitive skills should be used, knowing *how* to use them, and *why* to use them.

Metacognition is thinking about thinking. More precisely, we can think of metacognition as involving two central components (McCormick, 2003). The first is knowledge of cognition. Knowledge of cognition includes knowledge about cognitive

processing and barriers to it, how to use strategies to improve cognitive processing, when, and why these strategies are used. The second is regulation or control of cognition. This is the ability to implement, monitor, and revise these cognitive strategies to achieve one's learning goals.

For example, suppose that I am preparing to take an exam, and have to make decisions on how to study. My knowledge of cognition includes my knowledge about what study techniques are effective, the contexts in which those techniques are best used (e.g., a technique might be useful for one type of an exam, but not another, or one technique might be more useful than another given my current state of knowledge) and why these strategies are effective. I might recognize that simply re-reading the textbook will be ineffective, but that rewriting and reorganizing notes may be more effective. My ability to regulate my cognition will involve my actually implementing this knowledge and using the more effective technique, but also, checking to make sure that it is working. I might notice that I have achieved sufficient mastery of that particular content, or that the study skill is becoming less effective, and so revise the strategies I use to study. The learner who is not metacognitively aware may implement ineffective or inefficient study strategies, and the learner without regulative control might know what strategies are most appropriate, but fail to implement them.

The proposal then, is that critical thinking instruction involves teaching cognitive skills (the strategies) and the requisite knowledge of cognition (how, when and why to use these strategies) to empower students to regulate their own thinking. I will defend this proposal in two parts. First, I will argue that there are strategies that are effective at mitigating cognitive bias. Such strategies supplement the cognitive strategies already taught in critical thinking courses (such as argument diagramming). Second, I will argue that metacognitive knowledge and ability is required for students to implement these strategies effectively.

While the literature on cognitive bias is extensive, the literature on debiasing is much more limited (Lilienfeld et al., 2009). There are, however, a few strategies that are empirically supported (see Larrick, 2004 for a review). The *cognitive* debiasing strategy that is most well attested in the literature is "consider the opposite" or "consider an alternative" (Hirt & Markman, 1995; Lord, Lepper, & Preston, 1984; Mussweiler, Strack, & Pfeiffer, 2000). On this strategy, the reasoner is asked to generate reasons for the contrary position, or to explain an alternative outcome. Doing so directs the reasoners' attention to reasons for the alternative positions, rather than simply generating support-

ing reasons for the antecedently held position. This strategy has been found effective in reducing hindsight bias, as well as overconfidence bias (the tendency to be overconfident in the strength of one's own beliefs and judgments) and anchoring effects (the tendency to have one's judgments shaped by the initial piece of information received).

For example, Mussweiler et al. (2000), explored the effects of consider the opposite on real world judgments subject to anchoring effects. They provided car experts (mechanics and dealers, approached at their places of employment) with information on a car, including the sellers' estimate on the value of the car (the anchor). Mussweiler et al. manipulated two experimental conditions. The first condition was the value assigned to the car; some of the experts were given a high anchor (5000 German Marks) and others a low anchor (2800 Marks). The second condition was whether the participant was asked to provide arguments which were contrary to the anchor or not (consider the opposite). For example, participants using the consider the opposite strategy who received the high anchor were asked, "A friend of mine mentioned yesterday that he thought this value is too high. What would you say argues against this price?" (Mussweiler et al., 2000, p. 1145).

Among the participants who were not asked to employ a consider the opposite strategy, those who received the high anchor valued the car (on average) at 3563 Marks, against 2520 Marks for those who received the low anchor. This is a paradigmatic case of the anchoring effect, since the experts received the same information on the car, aside from the anchor. Among those who were asked to use a "consider the opposite" strategy, the anchoring effect was reduced (though not eliminated). Those who received the high anchor estimated the value of the car at 3130 Marks, versus 2783 from the low anchor/consider-the-opposite group. This reduced the high/low value spread from 29.3% to 11.1%. While the strategy did not eliminate the bias, it was effective in mitigating it.

This strategy shares common ground with the perspective taking strategy. In a series of papers, Galinsky and colleagues have found that having participants imagine what it would be like to be another person in a given scenario reduces intergroup bias (a bias in favor of members of one's own group and against outsiders) (Galinsky & Ku, 2004; Galinsky & Moskowitz, 2000; Todd & Galinsky, 2012; Todd, Bodenhausen, & Galinsky, 2012; Todd, Bodenhausen, Richeson, & Galinsky, 2011). In both the case of perspective taking and considering the opposite, the rea-

soner is tasked with making sense of another, rather than merely enumerating the reasons to support what s/he already believes.

Fernbach et al. looked at overconfidence in judgments about one's own understanding, and found that asking for causal, mechanistic explanations of how political policies work reduced this overconfidence, and mediated extreme political opinions on those policies (Fernbach, Rogers, Fox, & Sloman, 2013). Participants were asked for their opinion on various political policies (such as sanctions on Iran, single-payer health care, cap and trade), and to rate their level of understanding of that issue. Some of the participants were then asked to give a mechanistic explanation for how the policy works, while others were asked to justify their position. Afterwards, they were once again asked for their judgments on the policy, and on their own level of understanding. Those who gave an explanation expressed more moderate positions, and lower confidence, than those asked to justify their view. Like the consider the opposite and perspective taking strategies, asking the reasoner to explain or explore the issue prior to giving these reasons diminished the effects of cognitive bias.

There are also hints of other cognitive strategies available in the literature. Hafenbrack et al. found that practicing mindfulness through meditation reduced susceptibility to the sunk cost bias (the tendency to make judgments about future costs based on costs already paid) (Hafenbrack, Kinias, & Barsade, 2014). In one experiment, 78% of the participants led through a meditation session made judgments that indicate resistance to the bias, while only 44% of those in the control did so. While I will not pursue the pedagogical implications of mindfulness here, these results do suggest fruitful connections to be drawn with work in contemplative pedagogy (Barbezat & Bush, 2013; Hill, 2006).

Larrick (2004) also reviews a set of *social* strategies for debiasing. Groups, properly used, can diversify perspectives and diminish the impact of any single individual's biases. Models and decision support systems, which provide tools and explicit rules to guide decisions, also reduce the impact of bias by taking elements of the decision out of the hands of the reasoner.

While there is empirical evidence to support the effectiveness of these strategies, they are ineffective if not actually used when reasoning "in the wild." The metacognitive skill championed here is the ability and disposition to employ these strategies when appropriate. In such cases, the reasoner, rather than the instructor, is the trigger for the use of the effective debiasing strategies. That is, not only ought we prepare students on how to use debiasing strategies, but prepare them to be learners capable

of knowing when and how to implement them in actual reasoning scenarios. It is this more challenging demand which requires metacognitive skill.

To defend the role of metacognition here, it is worth pausing on the nature of these cognitive biases themselves. Prominent research programs in the psychology of reasoning, including Dual-Systems theory (see Kahneman, 2011 and Evans & Stanovich, 2013) and Gigerenzer's Adaptive Toolbox (Kruglanski & Gigerenzer, 2011), draw a distinction between quick, heuristic driven reasoning and deliberative reasoning. On Dual-Systems theory, these two types of reasoning are actually two distinct reasoning systems, while Gigerenzer draws the distinction based on the type of rules used in reasoning (satisficing versus optimizing rules). The first of the two reasoning systems on Dual-Reasoning theory, System 1, is quick, intuitive, automatic and driven by heuristics. The tasks undertaken by System 1 can vary in complexity, from orienting to the source of a sound, to driving a car (Kahneman, 2011, 21). One of the primary ways in which it operates is through heuristics, or procedures to determine adequate, though perhaps not correct, answers.² While these heuristics are useful in many cases, they can lead to error in other conditions.³

System 2, by contrast, is slow, deliberate, conscious, and effortful. Engaging this system is hard work, and we do not regularly do so. It is also key to overriding the heuristics and biases that originate in System 1 processing. System 1 processing, for example, often leads us to commit errors of statistical reasoning (such as base rate neglect, gambler's fallacy, and the conjunction fallacy). The tendency to commit these errors does not vanish when we develop our statistical knowledge. Rather, we become prone to overriding those initial judgments with the considered judgments of System 2. The cognitive debiasing strategies discussed above are tools for invoking System 2 processing to override the heuristic-driven responses of System 1.

² A heuristic need only be adequate because accuracy may come at too high of a cost. Evolutionarily speaking, it may be advantageous to decide quickly or with low effort rather than to take the time and energy to be correct. A simple example is predator identification—it is better to have a high number of false positives in identifying a nearby predator than it is to wait and ensure accuracy before acting.

³ It is worth noting, however, that though the cognitive biases typically arise out of System 1 processing, they do not exclusively do so. System 2 is also prone to bias and error, and the distinction between the two systems should not be read as the distinction between a normative (System 2) and non-normative (System 1) reasoning system. Rather, System 1 is well-suited for a range of tasks, but is error-prone in certain circumstances.

It is implausible to suppose that students (or any of us for that matter) will operate solely using System 2. Nor is it desirable. System 1 is not only less demanding, but it is highly effective in a wide range of situations. The aim, rather, is to help students develop the ability to invoke System 2 in the right conditions, in the conditions under which a bias is either likely, or likely to be costly (high stakes scenarios).

Wilson and Brekke's model of "mental contamination," or the unwanted influence of cognitive bias, details four decision points which determine whether our reasoning is contaminated by bias (Wilson & Brekke, 1994). First, is the subject aware of the bias? Second, is the subject motivated to correct it? Third, is the subject aware of the direction and magnitude of the bias? Fourth, is the subject able to adjust his or her response? If the answer to each is "yes," then mental contamination is avoided.

This model can be put in the vocabulary of metacognition. The first and third points concern *knowledge* of the cognition. They require that the subject have both general knowledge of cognitive bias, and the awareness of one's own cognitive processing to recognize the influence of the bias, or that one is in a situation particularly prone to the bias. The fourth is *regulation* of cognition. Does the subject have other strategies available to counteract the bias? On Wilson and Brekke's model, then, avoiding the effects of cognitive bias is a matter of recognizing a bias and implementing a response to minimize the effects of that bias. Doing so centrally involves metacognitive ability, both in terms of the knowledge of cognition, and the regulation of cognition.

The second decision point on Wilson and Brekke's model requires *motivation* to do something about the bias. The awareness of cognition and ability to regulate one's own cognition will be impotent if the reasoner does not desire to avoid the effects of the bias. This is consistent with findings on metacognition. Both metacognitive knowledge (Kraayenoord & Schneider, 1999) and metacognitive regulation (McWhaw & Abrami, 2001) have been found to be positively correlated with student interest. It also squares with the conceptions of the critical thinker championed by Ennis and Siegel, both of which emphasize the ways in which the critical thinker is not only someone who has the ability to think, but someone who actually does employ that ability, and is responsive to it.

The metacognitive approach to critical thinking pedagogy is designed to help students develop the knowledge of their own cognition required to recognize bias, and to develop a stable of strategies that can be implemented in the right conditions in or-

der to minimize its effects. I have argued for this in two parts. First, work in debiasing provides us with a set of strategies that have demonstrated effectiveness in mitigating the effects of cognitive bias. Second, on Wilson and Brekke's model, avoiding these biases requires metacognitive skill. The aim is to make the cognitive skills developed in a critical thinking course more regularly applied in new contexts beyond the end of term, making the skills taught in that course more durable and portable.

A metacognitive approach is not a panacea. Biases are potent, and they are widespread. What's more, the pedagogy is limited to the relatively narrow window of a course term. Being a critical thinker is a lifelong project. The metacognitive approach proposed here is not a solution to bias, but it is one way of mitigating the effects of bias by empowering students to respond to these biases in effective ways. Even if some biases prove to be recalcitrant, this model shows promise for helping students to avoid some and reduce the effects of others.

One might object that the use of deliberative reasoning to override heuristic driven reasoning sets the wrong goal. Our biases might be *features*, rather than bugs.⁴ Mercier and Sperber (2011) argue that the function of our reasoning abilities is primarily argumentative, rather than as a tool for rational thinking. We trade in arguments so that we can convince others to share our positions, and to produce better epistemic results from our collective reasoning. We are subject to confirmation bias because our goal is to convince others, and so we ought to make use of the evidence that best supports the claim we wish to defend. We are subject to motivated reasoning because we are proactively seeking arguments for anticipated dialogue with others.

Not only do these biases make us more effective communicators and persuaders (compare the interlocutor who is painstakingly careful about all sides and the one who forcefully makes the case for one position), but they also lead to better epistemic outcomes. There are two central reasons for this epistemic success. First, group discussions will have a variety of viewpoints that are each defended forcefully, and which the other members of that group have to confront. Consider the courtroom, where lawyers are tasked with representing particular perspectives to the best of their ability. The truth is not intended to emerge from the arguments of any single lawyer, but rather, from the process which brings together the strongest cases for the opposing views. Second, on Mercier and Sperber's model, these biases are less influential on us as *evaluators* of arguments (rather than as

⁴ I am grateful to an anonymous referee for raising this objection.

producers of arguments), at least when we are seeking truth rather than victory (Mercier & Sperber, 2011, 72).

If this is right, the objection goes, then why focus on methods that mitigate the effects of bias? These biases, it is contended, are central features of our reasoning abilities! The value of these biases, however, is at the group level, rather than at the individual level. If the goal of a critical thinking course is for students to be someone who is “appropriately moved by reasons” (Siegel, 1988), or someone with the right set of dispositions and abilities (Ennis, 1995), then our focus ought to be on improving the abilities of that individual student. Confirmation bias might help the individual contribute to the group, but at the expense of making that reasoner less likely to discern the truth him or herself. That is, the normative aims identified by Siegel and Ennis diverge from the evolutionary value that explains the argumentative model. We might still aim at Siegel and Ennis’s normative goals while acknowledging that the argumentative model better explains the function of reasoning.

If the normative aim was itself inconsistent with, or undermined, fulfilling this function, then the objection might persist. It might be that the function of reasoning is a normatively valuable goal for groups, and so we ought not undermine success at the group level in pursuit of success at the individual level. However, Mercier and Sperber themselves acknowledge that the debiased reasoner may be highly, and positively, influential on the group conversation. They do not warn us away from taking this as a normative goal, but only from assuming that such a reasoner is paradigmatic of a human reasoner (Mercier & Sperber, 2011, 73). It may well be that the individualistic conception of a critical thinker that I have adopted here, following Siegel and Ennis, is problematic and ought to be rejected in light of Mercier and Sperber’s model. Such questions, crucial as they are for guiding critical thinking pedagogy, are beyond the scope of this present project.

Finally, the seeking-truth attitude, rather than the seeking-victory attitude, is itself constitutive of being a critical thinker. That we are less prone to the biases when evaluating arguments with this attitude is no surprise—it is to say that the critical thinker is less prone to the biases. At issue on both views, then, is how to help our students take this attitude more readily, and the metacognitive strategies discussed here are intended to inculcate it. They do so by focusing our attention on understanding positions, rather than recruiting our “intuitive lawyer” (Baumeister & Newman, 1994) to defend our preconceptions.

Gerd Gigerenzer similarly challenges the classification of these tendencies as biases, noting that many heuristics are actually just as reliable as, and sometimes even more reliable than, optimizing rules which account for much more data. For example, investors following the simple “1/N Rule” (split one’s money evenly across N funds) perform just as well as investors using any of 12 approaches designed to estimate optimal allocation (Gigerenzer, 2008). Our heuristics, he argues, are in many cases, ecologically rational. That is, even if the rule is not logical in the traditional sense (that is, when evaluated independent of a context of application), it may produce equally accurate results when implemented in the right conditions.

This might be used to resurrect the challenge. Our heuristics are reliable when applied in actual contexts, and metacognitive awareness might impede our ability to deploy these heuristics. The ballplayer who focuses too much conscious attention on his or her swing, or on how to catch a flyball, often performs worse than when doing the task as a matter of routine (Kruglanski & Gigerenzer, 2011, p. 102). Might this approach similarly turn students into less effective reasoners, no longer able to use these valuable heuristics?

It is worth noting that the aim is not for students to constantly be using System 2, or deliberative reasoning. System 2 itself is prone to errors and ought not be held up as the normative standard. Rather, the aim is for students to utilize System 2 or deliberative reasoning when one’s heuristics/biases are liable to lead one into epistemic trouble. In investing, this might mean using the consider-the-opposite strategy when considering a fund that one has a vested interest in, but relying on a heuristic when more advanced statistical models do not offer any appreciable improvement.

Boudry et al. point out that “rationality...consists of using our reasoning abilities appropriately to deal with the situation at hand, not blindly following heuristics of which—with hindsight—we can appreciate the adaptive rationale” (2015, p. 531). We may appreciate that the gambler’s fallacy arises out of an ecologically rational process in a natural environment with limited probabilistically independent events, but nevertheless hold that the rational gambler is one who is able to recognize and overrule this heuristic. Or we may value the recognition heuristic when making judgments in ignorance, while being aware that advertisers can use this to judge brands based on their recognition factor (Boudry et al., 2015, pp. 530-531).

The critical thinker is able to implement the right tools for the job in a given situation. To some extent, the right tools re-

mains an open empirical question—we do not yet know which heuristics will be more accurate under what conditions. In the next section, I sketch out an approach that aims at the context-sensitive use of metacognitive response. While this approach is based on the assumption that biases like motivated reasoning, confirmation bias and overconfidence bias are problematic in high-stakes scenarios, the details of the approach could be altered as more empirical data comes in on the most effective uses of heuristics.

3. The implementation

This approach might be implemented in a wide variety of ways. Crucially, developing metacognitive skill involves a high amount of structured assignments designed to transition the learner from seeing a skilled practitioner model it to full and autonomous practice (Bruer, 1993). Here I will review one way to structure these assignments, based on work by Schraw (1998). Schraw introduces two tools, one designed to aid in the development of knowledge of cognition, the other in regulation of cognition. These tools can be adapted to tools across the curriculum and used to devise specific assignments to practice metacognitive skill.

To promote knowledge of cognition, he proposes using a “Strategy Evaluation Matrix.” This tool is a table of strategies that contains information on *how* the strategy is used, *when* it is used, and *why* it is used. Students develop their matrix over the course of the semester, building their skill in each strategy before introducing the next. As they work on a new strategy, students will be expected to make use of that strategy as well as all of its predecessors. The aim is that by the end of the semester, students will be proficient in each of the strategies, and more importantly, able to recognize the conditions underwhich the strategies should be implemented.

The contrast is with teaching students only the *how* (and perhaps the *why*) for the strategies. The *when* component is the crucial piece of information for metacognitive awareness. Students need to be monitoring the argumentative context for when particular strategies should be applied. Not only is the aim for students to be more proficient at recognizing the appropriate contexts for each strategy, but also to be regularly monitoring the situation for such contexts.

This awareness itself needs to be practiced. The Strategy Evaluation Matrix is a particularly useful tool for this, since it

can be used to structure assignments. Rather than giving students the vague and difficult direction to pay attention to their own reasoning, students can be tasked with doing specific assignments and reflecting on when they used different strategies, or to complete an assignment looking for an appropriate case to invoke a particular strategy. Used consistently throughout a semester, these structured assignments can slowly give way to increasingly open-ended instruction. Since the Strategy Evaluation Matrix guides the practice, students will have the resources to take responsibility for more and more of the monitoring as the semester wears on.

Schraw's matrix is designed for teaching reading skills, but the debiasing strategies discussed above provide the tools for a class on critical thinking. In particular, consider-the-opposite is a crucial strategy to include. The *how* direction can explain to students that they should try to put themselves in the position of someone who believes the opposing (or different) view, and to give the best case they can for it. To complete this reflection, one should be able to provide a satisfying (to the interlocutor) answer to the question, "Why might that person believe as they do?" The *why* condition is to be a charitable and intellectually virtuous reasoner, and also to communicate more effectively with people with opposing views.

The trigger, or *when* condition is subject to two pressures. On one hand, it ought to be responsive to empirical evidence about when deliberative reasoning is most effective, and when relying on heuristics is preferable. This will be particularly true for reasoning strategies where there is more evidence about the usefulness of heuristics, such as statistical strategies, and which might have more specific trigger conditions. On the other hand, it ought to be simple enough for students to remember and make use of in actual reasoning scenarios. To navigate between these two demands, I propose the following two-part condition:

Consider the Opposite—When Condition

- (a) when reasoning about high-stakes issues (e.g., issues which involve your strongly held values or have significant implications for yourself or others); and/or
- (b) when reasoning about issues about which there is significant disagreement.

These conditions do not perfectly capture cases where consider-the-opposite is valuable. The absence of significant disagreement need not mean there are no opposing arguments worth taking seriously (people might be wrong to agree so readily),

disagreement might not be reasonable (one position might be poorly defended) and considering the opposite can be valuable even for low-stakes reasoning. The aim, however, is to provide a condition which is broad enough for students to remember and apply “in the wild,” but which is also not so broad as to apply to all reasoning situations. Significant disagreement is used as a proxy for cases where there are available counter-arguments worth taking seriously. The disjunctive structure of the two conditions, and the use of actual disagreement in condition (b), is designed to respond to the false consensus bias, which here might lead people to assume that their position is more commonly believed than it really is. The focus on high-stakes scenarios in condition (a) is intended to limit the scope of the strategy with the aim of increasing its use by more clearly marking the situations in which it should be used. Identifying high-stakes scenarios is, of course, context-sensitive. In reflective exercises, students can be asked to identify which trigger condition applied, and why, so that they begin to reflect on what counts as a high-stakes scenario.

Additional strategies might include the use of argument mapping, and concept definition and explanation (both tasking students with explaining how something works, an understanding rather than justificatory task). To recognize situations where motivated reasoning is likely, I task students with doing an analysis of their own values when encountering new ideas. This task requires that students simply write down, for themselves, what their values are on the issue going into the text. The aim is to make salient the places where motivated reasoning might be particularly likely to occur, and what views the students might be hesitant to give up. This works together with the consider-the-opposite strategy, as a student might recognize that a particular reasoning context is high-stakes for them, given the way their values are implicated in that context. The matrix could also be supplemented with social debiasing strategies, such as structured group work to combat false consensus bias.

Schraw’s strategy for developing the regulation of cognition is the use of regulatory checklists. A regulatory checklist provides students with a set of questions to ask when planning, monitoring and evaluating their performance on a task. For example, a student might be asked to identify their goals before beginning (planning), to check whether they are making progress towards their goals (monitoring), and finally whether they achieved their goals (evaluating). The aim of the checklist is to get students in the habit of checking their performance and adjusting it by providing clear structure to guide the students. This

same task is accomplished by providing structured assignments on the Strategy Evaluation Matrix.

The intellectual virtues, or an operationalization of them such as Damer's (2008) "Code for Intellectual Conduct," provide a useful frame for a regulatory checklist. In both the monitoring and evaluating stages, students should ask questions like: "Have I considered the possibility that my own view might be false?" and "Would the opposing side agree with the way I have put their arguments?" Keeping a short (to encourage students to use it outside of class) list, tied into the normative and philosophical standards for critical thinking, provides the impetus for students to use the strategies. That is, the aim of the checklist is to encourage students to make use of the information they possess, and which is contained on the Strategy Evaluation Matrix, on how to avoid or mitigate the effects of cognitive bias. Finally, using a checklist follows from research on the advantages of an external checklist over memory or internalized routines for complex tasks (such as for pilots and nurses, usefully summarized in Gawande, 2011).

Schraw's tools provide a potent set of tools for implementing the metacognitive approach defended here. In particular, the Strategy Evaluation Matrix is an aid for students (in keeping track of the information they are expected to learn about the strategies) and for instructors (as the structure around which to build assignments to practice metacognitive skill). The students who improve their ability to use these debiasing strategies, and their tendency to actually use that ability, will improve their critical thinking skills by protecting themselves from the effects of cognitive bias. This, in turn, will help students develop into the intellectually virtuous and rigorous reasoners that critical thinking courses aim at, goals that are threatened by the potency and ubiquity of these biases.

This approach is intended to be adaptable to a wide range of pedagogical contexts, including different levels and subjects. The instructor can implement the strategy evaluation matrix using the strategies appropriate to their particular course, with assignments to fit those strategies. In running workshops on this approach, I have found that faculty across the curriculum consistently devise new and interesting strategies to apply this to their own field (from interpreting visual representations of quantitative data, to inhabiting a fictional character's socio-cultural position). Activities designed to encourage student motivation will similarly depend upon the pedagogical aims of your class and the elements of the critical thinker that the class focuses on. I will close with two general remarks about motivation, with

specific examples of how I implement them in my own course on critical thinking.

First, motivation is complex, and can be approached across a number of different levels (Brophy, 2013). One might use extrinsic rewards to motivate students to engage in particular actions that will help them master a skill or develop the habit of using a skill (e.g., giving credit for assignments completed). One might help students articulate and identify goals that they themselves affirm, and which motivate them to use the skills as a furthering of the goal. One might help students see the practice as intrinsically motivating. In a college classroom, one might work towards all at once, with different techniques used to develop each.

Second, motivation can be based on extrinsic and intrinsic factors—both on what rewards will motivate, but also on what the students are already motivated by. For example, one obvious source for motivation is from students' own lives and experiences. Critical thinking instructors often do this by choosing examples from contemporary politics and culture, examples which show the importance of clear thinking to our lives. This can also be taken further, into the smaller issues that students face on a regular basis.

For example, I task students with journaling exercises based around the strategy evaluation matrix where they are asked to report on a case where they applied one of the strategies to an argument they actually made in the past week. This low-stakes exercise is intended to not only cultivate the habit of metacognitive reflection, but for students to make clear to themselves each week how this sort of thinking benefits them (whether it be in making decisions about investments, or decisions about which campus eatery to patronize that evening). The grades students receive for doing the activity serves as a motivator of the first sort, while content of the activity is designed to help students articulate metacognitive awareness as a self-affirmed goal.⁵ This

⁵ This approach has much in common with literature on self-regulation (Boekaerts, Pintrich, & Zeidner, 2005). While precise characterizations of self-regulation are contentious, Pintrich's four phase model maps onto the approach defended here. Metacognitive knowledge is involved both in the forethought/planning phase, as well as the monitoring phase, where students articulate how, when, and why to use the strategies and then monitor the context for the situations in which their application is appropriate. Regulation of cognition occurs in the control phase. The Strategy Evaluation Matrix, and assignments based upon it, provides students with the opportunity for reaction and reflection. Assignments intended to foster motivation by helping students to connect their metacognitive reflection to their own values and goals are a form of *motivational monitoring*. Such motivation encourages the

content draws on the values and aims the students bring to the class, using their own pre-existing motivations. Critical thinking is a skill that can be tied into a variety of concerns and values that students bring to our classrooms, from its usefulness in achieving instrumental goals, to its role in the well-lived life. The motivations for pursuing and practicing this skill should similarly draw from across these concerns and values.

References

- Barbezat, D.P., & Bush, M. (Eds.). (2013). *Contemplative Practices in Higher Education: Powerful Methods to Transform Teaching and Learning*. San Francisco, CA: Jossey-Bass.
- Baumeister, R.F., & Newman, L. S. (1994). Self-Regulation of Cognitive Inference and Decision Processes. *Personality and Social Psychology Bulletin*, 20 (1): 3–19.
- Boekaerts, M., Pintrich, P.R., & Zeidner, M. (Eds.). (2005). *Handbook of Self-regulation*. Burlington, MA: Academic Press.
- Boudry, M., Vlerick, M., & McKay, R. (2015). Can evolution get us off the hook? Evaluating the ecological defense of human rationality. *Consciousness and Cognition*, 33: 524–535.
- Brophy, J.E. (2013). *Motivating Students to Learn*. Routledge.
- Bruer, J.T. (1993). *Schools for thought: a science of learning in the classroom*. Cambridge, Mass.: MIT Press.
- Damer, T. E. (2008). *Attacking Faulty Reasoning: A Practical Guide to Fallacy-Free Arguments* (6th edition.). Australia; Belmont, CA: Cengage Learning.
- Ennis, R.H. (1995). *Critical Thinking* (1st ed.). Upper Saddle River, NJ: Pearson.
- Evans, J.S.B.T., & Stanovich, K.E. (2013). Dual-Process Theories of Higher Cognition Advancing the Debate. *Perspectives on Psychological Science*, 8 (3): 223–241.
- Fernbach, P.M., Rogers, T., Fox, C.R., & Sloman, S.A. (2013). Political Extremism Is Supported by an Illusion of Understanding. *Psychological Science*, 0956797612464058.

student to continue implementing, and refining their implementation of, these strategies.

- Galinsky, A.D., & Ku, G. (2004). The Effects of Perspective-Taking on Prejudice: The Moderating Role of Self-Evaluation. *Personality and Social Psychology Bulletin*, 30 (5): 594–604.
- Galinsky, A.D., & Moskowitz, G.B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78 (4): 708–724.
- Gawande, A. (2011). *The Checklist Manifesto: How to Get Things Right* (Reprint edition.). New York: Picador.
- Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, 3 (1): 20–29.
- Hafenbrack, A.C., Kinias, Z., & Barsade, S.G. (2014). Debiassing the Mind Through Meditation Mindfulness and the Sunk-Cost Bias. *Psychological Science*, 25 (2): 369–376.
- Hill, C. (Ed.). (2006). *Special Issue on Contemplative Practices and Education*. Teacher's College Record (Vol. 108).
- Hirt, E.R., & Markman, K.D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, 69 (6): 1069–1086.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus; Giroux.
- Kahneman, D., & Riepe, M.W. (1998). Aspects of Investor Psychology. *The Journal of Portfolio Management*, 24 (4): 52–65.
- Kelemen, W.L., Frost, P.J., & Weaver, 3., C.A. (2000). Individual differences in metacognition: evidence against a general metacognitive ability. *Memory & Cognition*, 28 (1): 92–107.
- Kenyon, T., & Beaulac, G. (2014). Critical Thinking Education and Debiasing. *Informal Logic*, 34 (4): 341–363.
- Kraayenoord, C. E. van, & Schneider, W. E. (1999). Reading achievement, metacognition, reading self-concept and interest: A study of German students in grades 3 and 4. *European Journal of Psychology of Education*, 14 (3): 305–324.
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118 (1): 97–109.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108 (3): 480–498.
- Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making*. (pp. 316–337). Malden: Blackwell Publishing.

- Lilienfeld, S. O., Ammirati, R., Landfield, K., Nisbett, R., Ross, L., & Gilovich, T. (2009). Giving Debiasing Away: Can Psychological Research on Correcting Cognitive Errors. *Perspectives on Psychological Science*, 4 (4): 390–398.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47 (6): 1231–1243.
- Maynes, J. (2013). Thinking about Critical Thinking. *Teaching Philosophy*, 36 (4): 337–351.
- McCormick, C. (2003). Metacognition and learning. In W. Reynolds & G. Miller (Eds.), *Handbook of Psychology: Educational Psychology* (pp. 79–102). Hoboken, NJ: John Wiley; Sons.
- McWhaw, K., & Abrami, P. C. (2001). Student Goal Orientation and Interest: Effects on Students' Use of Self-Regulated Learning Strategies. *Contemporary educational psychology*, 26 (3): 311–329.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *The Behavioral and brain sciences*, 34 (2): 57–74; discussion: 74–111.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26 (9): 1142–1150.
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13 (3): 279–301.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26: 113–125.
- Schraw, G., Dunkle, M.E., Bendixen, L.D., & Roedel, T.D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology*, 87 (3): 433–444.
- Scott, B.M., & Berman, A.F. (2013). Examining the domain-specificity of metacognition using academic domains and task-specific individual differences. *Australian Journal of Educational & Developmental Psychology*, 13: 28–43.
- Siegel, H. (1988). *Educating Reason: Rationality, Critical Thinking and Education*. New York: Routledge.
- Stanovich, K.E., & West, R.F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of personality and social psychology*, 94 (4), 672–695.
- Thagard, P. (2011). Critical Thinking and Informal Logic : Neuropsychological Perspectives. *Informal Logic*, 31 (3): 152–170.

- Todd, A.R., & Galinsky, A.D. (2012). The reciprocal link between multiculturalism and perspective-taking: How ideological and self-regulatory approaches to managing diversity reinforce each other. *Journal of Experimental Social Psychology*, 48 (6): 1394–1398.
- Todd, A.R., Bodenhausen, G.V., & Galinsky, A.D. (2012). Perspective taking combats the denial of intergroup discrimination. *Journal of Experimental Social Psychology*, 48 (3): 738–745.
- Todd A.R., Bodenhausen, G.V., Richeson, J.A., & Galinsky, A.D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, 100 (6): 1027–1042.
- Veenman, M.V.J., Elshout, J.J., & Meijer, J. (1997). The generality vs domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction*, 7 (2): 187–209.
- West, R.F., Meserve, R.J., & Stanovich, K.E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of personality and social psychology*, 103 (3): 506–519.
- Willingham, D.T. (2007). Critical Thinking. *American Educator*, 31: 8–19.
- Wilson, T.D., & Brekke, N. (1994). Mental contamination and mental correction: unwanted influences on judgments and evaluations. *Psychological bulletin*, 116 (1): 117–142.