

Come Now, Let Us Reason Together: Cognitive Bias, Individualism, and Interactionism in Critical Thinking Education

AUSTIN DACEY

Department of Humanities

Mercy College

555 Broadway, Dobbs Ferry, NY, 10522, USA

adacey@mercy.edu

Abstract: In defending a new framework for incorporating metacognitive debiasing strategies into critical thinking education, Jeffrey Maynes (2015; 2017) draws on ecological rationality theory to argue that in felicitous environments, agents will achieve greater epistemic success by relying on heuristics rather than more ideally rational procedures. He considers a challenge presented by Mercier and Sperber's (2011; 2017) "interactionist" thesis that individual biases contribute to successful group reasoning. I argue that the challenge can be met without assuming an individualist ideal of the critical thinker as a solitary reasoner. Focusing on cognitive laziness and myside bias, I then argue that a more complete reckoning with the implications of interactionism about reasoning will require us to transcend individualism more fully to embrace the selection, design, regulation, and navigation of dialogic environments as central pedagogical aims of critical thinking education.

Résumé: En défendant un nouveau cadre pour incorporer des stratégies métacognitives qui éliminent les biais dans l'éducation à la pensée critique, Jeffrey Maynes (2015; 2017) s'appuie sur la théorie de la rationalité écologique pour affirmer que dans des environnements heureux, les agents obtiendront un plus grand succès épistémique en s'appuyant sur l'heuristique plutôt que sur des procédures plus rationnellement idéales. Il prend en considération un défi présenté par la thèse «interactionniste» de Mercier et Sperber (2011; 2017) selon laquelle les biais individuels contribuent au succès du raisonnement de groupe. Je soutiens qu'on peut relever le défi sans supposer un idéal individualiste du penseur critique en tant que raisonneur solitaire. En me concentrant sur la paresse cognitive et le biais de chacun, je soutiens ensuite qu'une évaluation plus complète des implications de l'interactionnisme sur le raisonnement nous obligera à transcender plus pleinement l'individualisme pour accepter la sélection, la conception, la régulation et la navigation des environnements dialogiques comme objectifs pédagogiques centraux de l'éducation à la pensée critique.

Keywords: cognitive bias, critical thinking, dialogic pedagogy, ecological rationality, individualism, interactionism

1. Introduction

Theorists and practitioners of critical thinking education are increasingly arguing that mitigating cognitive bias should be one of its proper aims (Bishop and Trout 2005; Kenyon 2008; Maynes 2015; Thagard 2011). At the same time, some are exploring instructional strategies that target not only the knowledge, skills, and dispositions of individual agents but also features of the physical, social, and institutional environments that agents inhabit and construct (Kenyon and Beaulac 2014, 2018). This development is welcome not just because of the encouraging empirical support enjoyed by such approaches to debiasing. It also goes some way to advancing the debate over long-standing critiques of the dominant critical thinking tradition as excessively focused on individual skills.

While debiasing projects typically draw on evidence from the “heuristics and biases” research program, an alternative research program has been put forward by Gert Gigerenzer and Peter Todd under the label *ecological rationality*. Ecological rationality rejects a “classical” conception of rationality modeled on general, formal, content-free statistical procedures such as Bayes’s Rule as the normative ideal, compared to which human reasoning emerges as a poor approximation (Goldstein and Gigerenzer 2002). Instead, it conceives of heuristics as an “adaptive toolbox” of “fast and frugal” decision procedures that have been shaped by natural selection to solve domain-specific informational problems where time, knowledge, and resources are limited (Gigerenzer and Todd 1999). Due to an evolved fit between these procedures and the “useful patterns of information available in the world,” they can be equally or more reliable as compared to effortful, reflective calculation (Todd and Gigerenzer 2007, p. 167). On this view, a heuristic that is less than individually rational in classical terms can be ecologically rational “to the degree that it is adapted to the structure of an environment” (Gigerenzer and Todd 1999, p. 19).

Consider the *recognition heuristic*: if one of two objects is recognized and the other is not, infer that the recognized object has a higher value with respect to the criterion in question (Goldstein and

Gigerenzer 2002). The recognition cue enables agents with little information about two cities, for instance, to judge which has a larger population more accurately than agents with more information because, owing to the greater availability of information about larger cities, recognizability is sufficiently correlated with population. Here, “less is more” because the inclusion of additional, less strongly correlated information would lead to a less accurate judgment. Recognition is rational in that it is reliably accurate given limited information; its rationality is *ecological*, rather than individual, in that the validity of the cue is neither internal to nor accessible by the agent.

The framework of ecological rationality presents a challenge to debiasing instruction that has been taken up by Jeffrey Maynes (2015, 2017). The concern is that by mitigating heuristic and “biased” strategies, such instruction may actually make agents worse off as reasoners: “If, then, critical thinking education dislodges the habit to use a strategy which is ecologically rational, and instead encourages the use of one which is slower and more error-prone, then that education may actually be harmful” (Maynes 2017, p. 120).

This risk is particularly salient in light of the “interactionist” perspective on reasoning advanced by Hugo Mercier and Dan Sperber (2011, 2017) according to which reasoning is an evolved adaptation with a social, communicative function: producing arguments to justify one’s positions and evaluating the arguments of others. Interactionism attempts to account for much of the evidence of purported failings of reasoning by arguing that these are in fact adaptive features of reasoning operating under evolutionarily non-normal experimental conditions of solitary racionation. It also predicts that in the more felicitous conditions of group dialogue, our reasoning capabilities would produce much better epistemic outcomes, a prediction borne out by considerable evidence surveyed below. Maynes glosses these reasoning features as “ecologically rational heuristics” (2017, p. 118) and observes that the pedagogical aim of mitigating them is in tension with the aim of preserving the value of their contributions to group reasoning.

For example, we are prone to confirmation bias because our aim is persuasion, and so we will cobble together the evidence that will

best convince our audience. This, in turn, improves the epistemic position of the group. Since many members of the group are forcefully defending their own viewpoints, the entire group has to confront a range of arguments. The truth need not emerge from the efforts of a single interlocutor, but rather it emerges from this group conversation. (p. 119)

The *interactionist challenge* to critical thinking education, then, is the risk that by increasing the quality of individual reasoning, we will diminish the quality of group reasoning, and by promoting the quality of group reasoning, we will undermine the pedagogical goal of mitigating bias in individual reasoning. While Maynes addresses this challenge, he also countenances the possibility that the individualism of the normative ideal of critical thinking may need to be reconsidered in light of Mercier and Sperber's work (Maynes 2015, p. 193).

In what follows, I examine the interactionist challenge and the response offered by Maynes as it is the only such attempt of which I am aware. I argue that the challenge, as he presents it, can be met without relying on an individualist ideal of the critical thinker as a solitary reasoner. Concentrating on cognitive laziness and myside bias, I then argue that a more complete reckoning with the implications of interactionism about reasoning will require us to transcend individualism more fully to embrace—along with individual metacognitive strategies of the sort defended by Maynes—the selection, design, regulation, and navigation of dialogic environments as central pedagogical aims of critical thinking education. The discussion concludes with a sketch and defense of a model that includes group-specific as well as individual-specific knowledge, skills, and dispositions.

Mercier and Sperber's interactionist view

The voluminous findings of the heuristics-and-biases literature would appear to deliver a dismal verdict on human reasoning abilities (Ariely 2008; Kahneman 2011). They also present a troubling anomaly for a commonsense understanding of the function of these abilities. According to this commonsense view, which Hugo Mercier and Dan Sperber (2017) call the *intellectualist view*, the purpose

of reasoning is “to enhance individual cognition” (p. 179); that is, “to help individuals draw better inferences, acquire greater knowledge, and make better decisions” (p. 182). By assessing intuitive beliefs and judgments, reasoning can detect and correct errors. Yet as the literature shows, this not what human beings typically do. From the perspective of the intellectualist view of reasoning, it would be profoundly surprising and perplexing if this central human faculty turned out to be so bad at performing its function (p. 179).

Beginning with an article in *Behavioral and Brain Sciences* (Mercier and Sperber 2011) defending an “argumentative theory of reasoning,” Mercier and Sperber have presented a radically different answer to the question of what reasoning is for. Their answer is located within an adaptationist framework. The primary adaptive function that reasoning was shaped by natural selection to perform is not intellectual or cognitive but social and communicative: to produce arguments to justify oneself to others and to evaluate the arguments of others. For highly social animals whose reproductive fitness was determined in large part by social coordination, conflict, and cooperation, the ability to produce reasons enabled communicators to persuade others to believe and do what the communicators desired whereas the ability to evaluate the reasons of others enabled audiences to exercise “epistemic vigilance”: benefiting from others’ information and judgments while simultaneously minimizing the chance of being misled by trusting in their testimony alone (2017, p. 233). The normal conditions for the use of reasoning, “the conditions to which they are adapted,” then, are social, and more precisely, *dialogic* (p.247). While our capabilities for producing and evaluating reasons can be enlisted into silent thinking to ourselves, they were designed for talking with others.

Given an interactionist metatheory, the purported weaknesses of reasoning are not failures of reasoning, but “features” (p. 235). Mercier and Sperber consider a range of results in the literature—including the classic Wason and ball-and-bat tasks—and argue that they are artifacts of the prevailing experimental paradigms. Either participants are engaging in solitary racionation, or they are engaging with interlocutors (the experimenters) who may attempt to elicit arguments but typically do not offer counterarguments of their own to which participants can respond. It is not adaptive to produce

reasons for our commitments when no one is demanding them nor to produce reasons that undermine our commitments. In such an environment, reason production naturally idles. Other phenomena, such as motivated reasoning and “moral hypocrisy,” are argued to be likely under solitary as well as dialogic conditions given the argumentative function of reasoning but unlikely in either given the intellectualist view (Mercier and Sperber 2011). What is anomalous for the intellectualist is entirely expected for the interactionist.

Additionally, interactionism predicts what would be difficult to explain given intellectualist assumptions: that when humans *reason together* rather than individually, performance improves dramatically. This prediction is supported by a wealth of evidence (discussed in Mercier and Sperber 2011 and Mercier, et al. 2017). When research participants solve the Wason selection task in groups rather than as individuals, they select the correct answer at rates of 80% rather than 10% (Moshman and Geil 1998). When they discuss “intellective tasks” such as the bat-and-ball problem, their success increases significantly (Trouche, Sander, and Mercier 2014). When children and adolescents engage in dialogue about issues with others, they produce fewer superficial or circular arguments for their positions and more counterarguments against them (as will be discussed in a later section).

The interactionist claim is not that being in a collective somehow eliminates our native laziness and myside bias but that groups, when suitably composed and structured, create a “cognitive division of labor” that puts them to most efficient use (Mercier and Sperber 2011, p. 65). First, the contributions of interlocutors indicate to the agent which of her claims they will not accept without argument, which arguments they will likely find persuasive, and which counterarguments they will produce. This prevents the agent from expending unnecessary cognitive effort on generating in advance the large number of arguments or counterarguments that may be pertinent but not valuable for persuasive purposes: “Instead of anticipating the interlocutor’s counterarguments, one can simply let the interlocutor provide them” (2017, p. 235). Second, myside bias leads agents to be most effective at producing arguments for their favored position, and, in an intellectually diverse group, these arguments are then subject to the evaluation of interlocutors engaging in myside

reasoning on behalf of their own positions. In such an environment, it pays to be miserly.

Another predicted cognitive benefit of the dialogic division of labor is that interlocutors' evaluations of an agent's reasons will be more accurate than her own, which are expected to remain biased in argumentative contexts. While Mercier and Sperber construe the effectiveness of group reasoning generally as "indirect" evidence for this prediction, they also point to some direct evidence. For example, Cowley and Byrne (2005) found that attention to falsifying instances increases when a hypothesis is believed to have been produced by someone else. One particularly striking result supporting the asymmetry between self-evaluation and other-evaluation is provided by Trouche and colleagues, who found that faulty arguments for solutions to cognitive tasks were detected more readily when participants thought they were someone else's, even though—by means of a choice blindness manipulation—they were in fact their own (Trouche et al. 2016). Those participants who were successfully led to believe that their arguments were someone else's rejected them more than half of the time. Importantly, they were more likely to reject incorrect answers than correct answers.

The superior performance of reasoning under dialogic conditions is not easily explained by supposing that the correct answer of one cognitively adept member simply cascades throughout the group without argument. The experiments show that groups instead converge on correct answers when members are convinced by good arguments that their initial answers were mistaken (for review, see Mercier and Sperber 2011). Nor is it the case that groups simply perform at the level of their best-performing member (Moshman and Geil 1998; Laughlin et al. 2002, 2003, 2006). Moreover, the quality of the arguments, and not the confidence of the person advancing them, appears to be the most important factor in accounting for group success on cognitive problems (Trouche, Sander, and Mercier 2014).

This brief survey of the interactionist perspective is not intended to show that it requires no additional empirical support, particularly regarding the superior performance of reasoning in argumentative contexts, but rather to justify my assumption that it is plausible

enough to merit serious consideration by critical thinking theorists and educators.¹

Maynes on the interactionist challenge

Drawing on Schraw's (1998) "Strategy Evaluation Matrix," Jeffrey Maynes (2015, 2017) develops metacognitive tools for guiding learners in determining when, how, and why to deploy debiasing strategies such as *consider the opposite*, which has been shown to reduce overconfidence bias and the anchoring effect (Arkes 1991; Mussweiler et al. 2000). He accepts the thesis of ecological rationality that in certain environmental conditions, fast and frugal heuristics may be more epistemically successful than classically "ideal" procedures of effortful, maximizing calculation. Accordingly, his approach aims at cultivating in the learner a "metacognitive awareness of the conditions under which she ought to use ideal strategies, and the conditions under which the ideal strategies are not useful, or worse, harmful" (2017, p. 124).

Maynes considers an objection to his project that he derives from the interactionist perspective. Briefly presenting the interactionist argument that myside bias and cognitive laziness "lead to better epistemic outcomes" in group argumentative contexts because they are adapted to such contexts, he voices the concern that bias mitigation would therefore be "the wrong goal" for a pedagogy of critical thinking (2015, p. 192). Maynes (2015) responds by distinguishing between epistemic benefits to groups and epistemic benefits to individuals.

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 (p. 193).

¹ I am grateful to an anonymous reviewer for suggesting this qualification.

Maynes claims that while the argumentative model may provide a better explanation for the design and functioning of reasoning than the intellectualist model, this is not a basis for abandoning the improvement of individual reasoning as a *normative ideal* in critical thinking education. He goes on to formulate a normative extension of the objection: “It might be that the function of reasoning is a normatively valuable goal for groups, and so we ought not to undermine success at the group level in pursuit of success at the individual level” (Maynes 2015, p. 193). Yet, he points out, Mercier and Sperber themselves “acknowledge that the debiased reasoner may be highly, and positively, influential on the group conversation” (p. 193). They do not warn us against adopting individual debiasing as a normative goal, “but only from assuming that such a reasoner is paradigmatic of a human reasoner” (p. 193).

As I read Mercier and Sperber, their claim here is somewhat weaker than Maynes takes it to be. In the passage he cites, they do assert that an ability to anticipate counterarguments may be necessary “to excel” in groups, but they do not seem to be referring to epistemic excellence as such. They go on to mention medieval *disputationes* and scientific discourse as examples of this “valuable culturally acquired skill,” and conclude: “In most discussions, rather than looking for flaws in our own arguments, it is easier to let the other person find them and only then adjust our arguments, if necessary” (Mercier and Sperber 2011, p. 73). In a 2017 paper entitled “Natural-Born Arguers: Teaching How to Make the Best of Our Reasoning Abilities,” Mercier and colleagues take up at some length the pedagogical question of how to improve people’s reasoning in argumentative contexts (Mercier et al. 2017). While they consider group size, structure, conversational practice, and other factors likely to increase the availability of diverse arguments within a group, they say nothing about the importance of mitigating individual myside bias.

Setting aside this exegetical question, what are we to make of Maynes’ claim that critical thinking education should aim at “improving the abilities of that individual student” even when doing so may undermine values that accrue at the “group level”? There are two relevant ways to interpret this first phrase. It could mean improving an agent’s abilities *as exercised in solitude*, or it could mean

improving an agent's abilities *as exercised in dialogic settings*. According to the interactionist, myside bias makes an agent less epistemically successful than she would be without it when reasoning alone, but it makes an agent more epistemically successful than she would be without it when reasoning with others. On this view, mitigating myside bias may not be an effective strategy for improving an agent's abilities under dialogic conditions, even while it may be an effective strategy for improvement under solitary conditions.

In what sense is an agent better off individually by engaging in myside thinking in a dialogic context? It is not necessarily or conceptually true that an epistemic benefit to a group is a benefit to each of its members.² For example, a group may produce a solution to a cognitive problem that is superior to any individual solution but is not shared with individual members—imagine that the crowd at the country fair submits independent estimates of the weight of an ox, and the remarkably accurate result of averaging their (individually inaccurate) estimates is never revealed to them (Surowiecki 2004). Nevertheless, I know of no reason to expect that such circumstances will be typical of group reasoning in general. It could be that when a group solves a cognitive problem through argumentation, that solution and the argument for it are accessible to all individual members, simultaneously becoming their solution as well. Such is the case in the research on collective performance on Wason tasks, for instance. Here we find *agent success via group success*. In this way, an individual agent's myside contributions to group reasoning could be improvements to that agent's epistemic success relative to solitary reasoning on the (interactionist) assumptions that such contributions increase the likelihood of agent success via group success and that agent success via group success is more likely than agent success via solitary reasoning, even following debiasing training.

When we look beyond the experimental literature on group cognitive performance to consider group reasoning in the wild, the question is to what extent individual debiasing will be instrumental to agent success via group success. We should not presuppose that

² I will use “group” to encompass dyads: any set of one or more other persons with whom an agent has a verbal interaction characterized by some degree of interpersonal accountability (for more on accountability conditions, see below).

unbiased contributions to group reasoning will typically be more effective, nor can we presuppose that biased contributions will typically be more effective. As I suggest below, these outcomes will depend sensitively on contingent factors such as the cognitive diversity of the group. What is needed is a nuanced and empirically-grounded account specifying the contexts in which and the degree to which critical reasoners should rely on heuristics and biases in dialogic environments, and, extending Maynes' approach, what cues will most reliably guide their deployment.

Thus far I have argued that interactionism does not force us to choose between epistemic benefits to groups and epistemic benefits to the agents who are its members. In this respect, I disagree with Maynes that interactionism is for this reason in tension with an individualist normative ideal of the critical thinker as a solitary reasoner—the assumption that the knowledge, skills, and dispositions that constitute the normative critical thinker are those an agent would have in non-dialogic conditions. I will now argue that this ideal is nevertheless challenged in another way by interactionism about reasoning and by ecological rationality more generally. To that end, it will be helpful to clarify some varieties of individualism in the theory and pedagogy of critical thinking.

Two forms of individualism about critical thinking

One way to be an individualist about critical thinking is to hold that the knowledge, skills, and dispositions of critical thinking are best cultivated through pedagogical methods directed at individual learners. This is pedagogical individualism, of which we can distinguish between stronger and weaker versions. Strong pedagogical individualism would assert that critical thinking is best cultivated through individual learning rather than collaborative learning.

Strong pedagogical individualism (SPI): The most effective instructional strategies for cultivating critical thinking will be individual rather than collaborative or dialogic.

As an empirical thesis, SPI is *prima facie* implausible. Not only does it run contrary to the broad empirical support for collaborative learning generally (Johnson, Johnson, and Smith 2014) and a

longstanding strand of the dominant critical thinking tradition emphasizing the pedagogical value of dialogue and debate (Lipman 1991, 2003; Paul 1987; Sharp 1987, 2014). It is also undermined by a substantial body of literature, explored below, on the development of argument literacy in young people. Finally, while meta-analyses of the effectiveness of particular instructional strategies for supporting student learning of critical thinking have produced mixed and uneven results due to a lack of genuinely experimental designs in the literature (Behar-Horenstein and Niu 2011; Tiruneh, Verburch, and Elen 2014), the more rigorous studies point to the effectiveness of dialogic strategies such as small-group discussion (Abrami et al. 2015). The available evidence does not support a strongly individualist pedagogy. This should not be surprising given an interactionist metatheory about the design and proper function of human reasoning capabilities.

One could reject SPI, however, while embracing a weaker and more plausible individualism. On this view, while dialogic methods of instruction may produce better pedagogical outcomes, they do so not by targeting the features of groups or other aspects of the learners' environments but by targeting features of the individual learners themselves.

Weak pedagogical individualism (WPI): The most effective strategies for cultivating critical thinking will be designed to modify the intrinsic properties of individual agents, not their extrinsic properties or the properties of their physical, social, or institutional environments.

Sidestepping some philosophical brambles, we can characterize “intrinsic properties” roughly as interior properties—those that an entity has in virtue of how it is, how its parts are, and how those parts are related to each other—rather than in virtue of how it and its parts are related to other entities (Marshall and Weatherson 2018). On this understanding, *having mass* is an intrinsic property while *being an aunt* is a non-intrinsic or extrinsic property. To appreciate the content of WPI, consider the instructional strategy of training learners in the use of *consider the opposite* as a metacognitive technique for mitigating individual bias. One might hold that this technique is best cultivated by collaborative or dialogic learning even while

holding that the technique is (weakly pedagogically) individualistic. It is designed to impart to the agent some declarative knowledge of the technique and its importance and to develop in her the skill and the disposition to apply it in the appropriate contexts, ideally with the result of mitigating the bias of her reasoning. Here, the intended modifications to the relevant knowledge, skills, dispositions, and reasoning would all be modifications to the intrinsic properties of the agent.

Notice that even debiasing projects situated in a framework of ecological rationality could be weakly pedagogically individualistic. This is so despite the fact that the features that make a reasoning procedure ecologically rational include properties external to the agent. Consider again the recognition heuristic. Part of what explains the reliability of this heuristic, under ecologically felicitous conditions, is the strength of the objective correlation between the cue of recognition and the criterion in question, and this is a property that is not internal to or introspectively accessible by the agent. Nevertheless, a pedagogy that includes recognition among its metacognitive strategies could be weakly individualist inasmuch as it does not aim at modifying these external ecological features. It only aims at equipping agents to know how and why to deploy such strategies when ecological circumstances make a frugal, satisficing solution more optimal than a costly, maximizing one. Again, the knowledge, skills, and dispositions involved will all turn out to be constituted by some constellation of intrinsic agential properties.

Debiasing programs that aim at modifying an agent's environment, however, do challenge WPI. Kenyon and Beaulac's (2014, 2018) recent contributions to the debiasing literature serve to illustrate one such promising approach. They claim that the cognitive and social psychological literature demonstrates that at least for a wide class of biases, any debiasing strategy "intended to be learned and subsequently self-deployed by individuals, acting alone and at the point of making a judgment, is unlikely to succeed in significantly minimizing biases" (2014, p. 343). This is in part because it is "typical of biases that their distorting effects are invisible to the agent displaying them, at the point of judgment or action" (Beaulac

and Kenyon 2018, 94; see also Arkes 1981).³ Without denying the importance of agential approaches, they make the case for approaches that engineer the built and social environment to cue bias-mitigation efforts by individual agents at the moment of judgment or “nudge” their behavior towards more favorable outcomes even in the absence of attempted or successful agential debiasing.

Refining proposals by Larrick (2004) and by Soll and colleagues (Soll, Milkman, and Payne 2015) to distinguish between debiasing strategies that modify the agent and strategies that modify the environment, Beaulac and Kenyon (2018) construct a four-level taxonomy. The first two levels are directed at agents. Level 1 strategies are designed to reduce the dispositions of agents to produce biased judgments, and Level 2 strategies train agents to deploy cognitive techniques for mitigating biased judgments when they do occur. The aim at Level 3 is “[t]raining agents (individually or collectively) both to create and to defer to situational ‘nudges’ that debias otherwise distorted judgments in context,” and at Level 4, “[t]raining agents (individually or collectively) to create and defer to processes or other situational constraints that debias actions or outcomes even when individual judgments are distorted and uncorrected” (p. 95). Beaulac and Kenyon illustrate their taxonomy with the case of gender bias operating within a hiring committee. An example of a Level 3 intervention is priming committee members prior to deliberations with a presentation on biases in hiring and leaving a visual reminder of this presentation on the conference table with the effect of mitigating the bias of individual members. In a Level 4 strategy, an individual committee member may retain “an uncorrected bias of judgment against women in the profession; but anonymized applications hide candidates’ gender information, and the committee member ultimately (unknowingly) votes to hire a superior woman candidate” (2014, pp. 351-352).

Kenyon and Beaulac (2014) argue that the most promising programs of debiasing will engage all four levels, and thus that critical

³ Maynes’ approach attempts to avoid this problem by constructing metacognitive debiasing strategies that are deployed in response to an environmental cue, not the agent’s introspective awareness of bias. I am grateful to Jeffrey Maynes for this emphasis and for additional comments on a draft of this essay.

thinking education should be expanded to encompass, along with agent-level interventions, extensive “practical guidance on how to structure and engage with one’s environment to promote good reasoning” (p. 360). This will include “teaching how and why to adopt decision-making policies and evidence-gathering practices that do not require the virtuoso ability to rise above invisible and subtle biases” (p. 360) Thus, their ecological account of debiasing rejects WPI in that it asserts that at least some important instructional strategies in a critical thinking curriculum will be designed to indirectly or directly modify properties of agents’ physical, social, or institutional environments in addition to the intrinsic properties of individual agents.

The interactionist challenge reframed

Two features of reasoning that are central to the case for interactionism, cognitive laziness and myside thinking, map onto two canonical components of critical thinking in the dominant tradition: the ability and disposition to produce arguments for one’s commitments and the ability and disposition to produce counterarguments to those commitments. How can people develop these abilities and dispositions in the face of their natural tendencies to laziness and myside thinking? Interactionism suggests that the answer lies in argumentative dialogue (Mercier et al. 2017).

Myside bias is known to be pervasive and particularly resistant to debiasing efforts (Lilienfeld, Ammirati, and Landfield 2009).⁴ It does not appear to be correlated with intelligence, cognitive ability, or verbal ability (Stanovich, West, and Toplak 2013; Perkins 1985). Neither does it appear to be significantly reduced by direct instruction in formal logical rules or cognitive biases (Lehman and Nisbett 1990; Lilienfeld, Ammirati, and Landfield 2009). The most promising individual debiasing technique appears to be the metacognitive strategy *consider the opposite* (Lord, Lepper, and Preston 1984; Hoch 1985). From the perspective of interactionism, the relative

⁴ I here avoid the question of the relationship between myside bias and confirmation bias, which is the tendency to seek or interpret evidence “in ways that are partial to existing beliefs, expectations, or a hypothesis in hand” (Nickerson 1998, p. 175). For discussion, see Mercier (2017).

effectiveness of this strategy should not be surprising as it approximates individually and monovocally the experience of encountering contrasting and opposing voices in dialogue with others. Morris' (2017) recent review of the effects of oral argumentation on myside bias concluded that "[e]nabling argumentation among students minimizes myside bias more effectively than explicit instruction about argument forms or reminders about incorporating counterarguments" (p. 10).

A substantial body of research in educational and developmental psychology within a broadly Vygotskyian tradition reveals that children and adolescents who engage in oral argumentation with peers show increased competence in argument skills, including producing relevant arguments and considering counterarguments (Anderson et al. 1997; Anderson, Chinn, Waggoner, and Nguyen 1998; Anderson et al. 2001; Chinn and Anderson 1988; Kuhn, Shaw, and Felton 1997; Lao and Kuhn 2002). There is also growing evidence that gains in argumentative competence can transfer to other reasoning tasks such as writing persuasive essays on unrelated subjects. For example, students who participate in instructor-facilitated and peer-driven "collaborative reasoning" subsequently write essays containing more arguments and counterarguments than a control condition (Reznitskaya et al. 2001). Gains in argument skills—particularly considering and rebutting opposing views—are greater among those who engage in discussion than among those who receive direct instruction in argument skills alone (Reznitskaya, Anderson, and Kuo 2007; Kuhn and Udell 2003).

The mechanisms that might underlie this phenomenon are not entirely clear. One theoretical explanation lies in the cognitive division of labor proposed by Mercier and Sperber. Another is provided by research on the dialogic condition of accountability to others: "the implicit or explicit expectation of decision makers that they may be called upon to justify their beliefs, feelings, or actions to others" (Lerner and Tetlock 1994, p. 1). Accountability has been shown to influence the quality of reasoning by motivating agents to secure the approval of their audience. Particularly when the views of an audience are unknown and the agent has not publicly committed to a position, accountability increases the production of more complex arguments, the anticipation of criticisms, and the

generation of counterarguments (Tetlock, Skitka, and Boettger 1989). To account for transfer to new content domains and communication modalities, some argumentation researchers have invoked the construct of argument *schema*: content-neutral patterns of social, rhetorical, and formal dialogic moves that are acquired inductively and applied in new domains (Reznitskaya and Anderson 2002; Walton, Reed, and Macagno, 2008; for discussion, see Nussbaum and Asterhan 2016).

Taken together, the evidence recommends participation in argumentative dialogue as an effective strategy for addressing myside bias and cognitive laziness in both collective and individual reasoning conditions. Kenyon and Beaulac's four-part taxonomy provides a helpful way to frame this recommendation. Arguably, the limit case of "modifying" one's environment is *selecting* the environment one is in. If so, then an instructional strategy that trains agents to seek out a dialogic context for reasoning about certain kinds of decisions could meaningfully be described as being intended to modify agents' environments (in addition to agents). Insofar as seeking argumentation with others is part of a strategy for mitigating occurrent bias, such a technique could be located at Level 3. Insofar as it is part of a strategy for optimizing cognitive performance without mitigating an individual bias (in keeping with Mercier and Sperber's analysis of myside bias in argumentative contexts), such a technique could be located at Level 4.

Looking beyond this minimal sense of environment modification, we can recognize that dialogic environments are subject to continual influence by the behaviors of individual members. Further, dialogic environments are often influenced by structural features—such as cultural norms or technological nudges—that may be outside of members' immediate awareness or control. While interactionism about reasoning may superficially appear to present group argumentation as such a panacea, in fact it counsels careful attention to the features of groups that promote epistemically favorable outcomes, features which are far from given or guaranteed.⁵

⁵ I am indebted to K. Patrick Fazioli for prompting me to devote more attention to the potential cognitive failures of group reasoning. My views on critical

Take the ability to produce reasons for one's beliefs, for example. On an interactionist view, this ability is expected to be influenced by the cognitive availability of arguments in one's dialogic environment, which is, in turn, linked to the ease with which interlocutors can make them public (Mercier et al. 2017). If a group's conduct imposes reputational pressures that are too great, public reasons may be subject to rigorous evaluation by others, but interlocutors may be inhibited from making them public in the first place, thus reducing the availability of arguments (Mercier et al. 2017). Relatedly, evidence suggests that the average "social sensitivity" of group members and the "equality in distribution of conversational turn-taking" in a group is more strongly correlated with the group's performance on a variety of cognitive tasks than with the average or maximum 'individual intelligence factor' of members (Woolley, et al. 2010). But neither the level of reputational pressure nor the extent of conversational turn-taking in a group is constituted by the intrinsic properties of any one agent who belongs to it. You may be highly disposed to express an optimal degree of public sanction and to take turns conversationally, but if you find yourself in a leader-dominated, status-driven, or adversarial group, argument production, and thus your cognitive performance via the group, is predicted to suffer. One divergent swallow does not a spring make.

Or consider the ability to produce counterarguments to one's own views. As previously mentioned, accountability research demonstrates that the effects of accountability vary considerably depending on, among other factors, the intellectual composition of the audience to which a reasoner may have to provide justification. Agents who are accountable to an audience with a mix of views are better at anticipating reasons against their own views than are agents who are accountable to an audience that is homogeneously aligned with or opposed to them (Tetlock, Stitka, and Boettger 1989). In the former case, agents tend to rely on an "acceptability heuristic": telling people what they are likely to accept. In the latter, they tend to engage in "defensive bolstering": reinforcing their existing beliefs

thinking and dialogic pedagogy have been deeply influenced by our many conversations and collaborations.

to the neglect of countervailing reasons. Additionally, the composition of a group affects the ecological rationality of individual myside bias. If a group lacks sufficient intellectual diversity, myside thinking may lead to “informational cascades” (Sunstein and Hastie 2008) as all parties seek confirmation for the same beliefs. But the composition of a group is not an intrinsic property of any one of its members. Two agents with identical internal properties could occupy groups with widely differing degrees of intellectual heterogeneity. Nevertheless, the agent situated in the more heterogeneous group will probably reason better. Thus, in the case of two canonical critical thinking skills—producing arguments to support one’s views and producing counterarguments—carefully calibrated environmental modification and maintenance may be required to achieve optimal outcomes in group contexts.

In sum, interactionism posits that our reasoning capabilities yield better outcomes when exercised in dialogic conditions. It also draws attention to the ways in which reasoning performance in dialogic conditions depends on the design, maintenance, and navigation of those conditions. Taken together, these tenets suggest that critical thinking education should include extensive instruction in why and when to enter into dialogic environments and how best to promote good reasoning within them. In some cases, this instruction would impart strategies that mitigate individual reasoning patterns regarded as classically irrational, and in others, it would impart (Level 4-type) strategies that preserve them in the interest of ecological rationality—counseling agent success via group success. These strategies would include specifically dialogic knowledge, skills, and dispositions and would transcend weak pedagogical individualism by asking agents to aim at modifying social-ecological characteristics along with agential characteristics.

In this regard, an interactionist perspective converges with the longstanding charge against the dominant critical thinking movement that it has not devoted due attention to ways in which individual thinking is interwoven with social relations. For critical pedagogy theorists, “the object of thinking critically is not only against demonstrably false beliefs, but also those that are misleading, partisan, or implicated in the preservation of an unjust status quo” (Burbules and Berk 1999, p. 49). Crucially, this status quo encompasses

aspects of the educational environment itself, from the assumptions, norms, and values of its prevailing culture and curricula, both overt and “hidden,” to the “pedagogical relations” between instructors, learners, and peers (Apple 1971; Friere 1970, 1973; Giroux 1988). For pragmatist and feminist theorists such as Barbara Thayer-Bacon (2000), the most crucial social relations are those that exist between co-constructors of knowledge within a community of inquiry. Thus, “communication and relational skills” are central to what she calls “constructive thinking” (p. 165). Similarly, Ann Margaret Sharp, a pioneer in the “philosophy for children” (or philosophy with children) movement, conceives of a community of inquiry as being constituted by attitudes of care, including “care for the form of the dialogue” (Sharp 2014, 20).

In defense of weak pedagogical individualism, it might be objected that every ecological modification can be construed, and is more readily and usefully construed, as a corresponding modification to some intrinsic agential properties. In the case of regulating reputational pressures and maximizing argument availability, for instance, it could be maintained that the relevant environmental modifications will always be accomplished by individual agents following procedures by which they *modify themselves*: making sure that they apply the right degree of sanction to bad arguments, taking steps to make sure that others do as well, deferring to structural nudges that reinforce these practices, and so forth. Thereby, group-specific procedures would turn out to be individual properties after all.

It may be true that every group-specific procedure could be classified under some agent-specific procedure, but that in itself is not a compelling reason to do so. The ideal of the critical thinker is a model, and like all models—most paradigmatically, maps of terrains—it is best constructed at a scale of granularity that strikes the right balance between precision and comprehensiveness on the one hand and usefulness on the other. With too little granularity, the model omits information necessary to guide users. With too much, it includes information that overburdens or distracts users from what ought to be most salient for use. Indeed, any of the canonical critical thinking skills could be omitted from a model by assimilating it to a taxonomy that is less granular. For example, the task of *identifying*

assumptions could be assimilated to the less granular task description of *analyzing arguments*, which itself could be assimilated to the less granular task of *accepting arguments in proportion to their strength*, and so on, perhaps until we arrive at the coarsest grain: *being appropriately moved by reasons*. Why include any procedures that go beyond the ultimate criterion of rationality? Yet if we do include *seek counterarguments*, why not include *seek a heterogeneous dialogic context*?

The principled rationale for selecting a model at one scale of granularity over another is the likelihood that it will produce better outcomes, estimated on the basis of the available evidence about how users will use (and misuse) it. By this rationale, we have no non-question begging reason on conceptual grounds to restrict a model of critical thinking to those and only those features that can be described in terms of the intrinsic properties of individual agents. It will not do to insist that every model of critical thinking exclude features beyond those of a solitary reasoner on the grounds that the ideal of critical thinking is, conceptually, an ideal of solitary reasoners since this would simply beg the question against ecological rationalists and interactionists. We would have to ask whether the most practically valuable scale of granularity would include non-agential features. This is of course an empirical question.

One general reason to expect that the most practically valuable taxonomy will include non-agential features is that some group-level strategies are *interdependent*; their effectiveness as deployed by one member depends on the dispositions and behaviors of others. In this respect, they are characterized by a “collective action” dynamic. I suggested above that the epistemic value of an individual group member’s *myside thinking* is dependent on the presence of intellectually heterogeneous others who are disposed to argue vigorously for opposing positions. In another type of case, an agent who can consistently mitigate a cognitive bias through a Level 2-type strategy may have no rational need, considered in isolation, to adopt a Level 4-type solution. Nevertheless, adopting a Level 4 solution among the group could be rational for her group, and thus for her, given that others will remain subject to that bias. Intuitively, in such a case the most useful way for the agent to frame her intention is in the form, *we should adopt or defer to this strategy*, not *I should*

adopt or defer to this strategy. Here, the content of the strategy is group-specific, not agent-specific.

Another general reason to frame ecological modifications in group-specific terms is the efficacy of *roles* within groups, where these roles are described functionally in a way that does not refer to any one agent. Sunstein and Hastie (2014) emphasize the value to group deliberation of assigning roles such as “red team”—members tasked with critically examining group’s assumptions and disrupting group consensus by producing dissenting reasons. They also argue that in the interest of managing reputational pressures, those in the role of “leader” have a special reason to self-silence, particularly in the initial stages of deliberation, that does not apply to non-leaders. By definition, such roles are not occupied by all members of a collective but could in principle be occupied by any member. Therefore, strategies that concern role assignment or role conduct will be most usefully framed in functional terms and addressed to all; for instance, *we should initially silence the leader.* Here again, the content of the strategy will be group-specific, not agent-specific.⁶

Note that what is at stake here is more than a mere classificatory or explanatory difference. An account of critical thinking that fully embraces dialogic competencies will also make a *normative* difference—a difference in our assessment of the rationality or epistemic virtue of individual agents. An agent could contribute to an epistemically successful outcome for a group, and thereby herself, by engaging in myside and other motivated reasoning that would in solitary conditions be considered suboptimal. This is the kind of case that concerned Maynes. Or via a group interaction, an agent could arrive at an accurate judgment that is not grounded in sufficient reasons that are internal to and introspectively accessible by that agent. When accurate estimates are obtained by equally weighting and averaging the incorrect individual estimates across a group, for example, the internal reasons of any one agent are insufficient to ground an accurate judgment (Mannes, Larrick, and Soll 2012). In either

⁶ I cannot explore here an obvious connection to the substantial literature on “socially extended cognition” and “shared” or “we-intentionality” (Merritt and Varga 2013; Tomasello 2009). I owe this clarification to an anonymous reviewer.

case, from an interactionist perspective, we would be evaluating the agent's epistemic performance favorably despite the fact that it would be evaluated unfavorably as an instance of solitary reasoning. We would be ascribing rationality to an agent who lacks access to internal grounds that would make it possible, in Maynes' words, "to discern the truth him or herself." This implication of interactionism about reasoning intersects with debates over the compatibility of ecological rationality with internalist conceptions of rationality (Boudry, Vlerick, and McKay 2015) and epistemic "credit" as a condition for knowledge (Axtell 2017).

Conclusion

The implications of Mercier and Sperber's interactionist view of human reasoning for the dominant tradition of critical thinking education remain largely unexplored. One possible implication examined by Jeffrey Maynes is that efforts to mitigate myside-biased reasoning will be counterproductive in group contexts, where it is expected on the interactionist view to contribute to good epistemic outcomes. I have suggested that the apparent tension between group-level outcomes and agent-level outcomes can be addressed by recognizing the phenomenon of agent success via group success. Typically, when a group reasons well together, its members are epistemic beneficiaries. Therefore, by maintaining that myside-biased reasoning can be ecologically rational in felicitous dialogic contexts, we are not thereby committed to prioritizing the value of group success over the value of agent success or "collective" rationality over individual rationality.

Nevertheless, interactionism supports the case that dialogic argumentation is a particularly promising strategy for managing myside thinking and cognitive laziness, both as a means of reasoning well with others when these patterns are in effect and as a means of mitigating their effects when reasoning alone. In this way, it challenges an individualist assumption within the critical thinking tradition that I have called weak pedagogical individualism: the thesis that the most effective instructional strategies for cultivating critical thinking will be designed to modify the intrinsic properties of

individual agents, not their extrinsic properties or the properties of their physical, social, or institutional environments.

Against this assumption, I have argued that participation in critical dialogue can be an effective strategy for addressing laziness and myside thinking and that the selection, design, regulation, and navigation of felicitous dialogic environments will be facilitated by a repertoire of distinctively dialogic knowledge, skills, and dispositions. This repertoire would transcend individualism by framing dialogic strategies in group-specific terms and evaluating agents who conform to them as epistemically successful even when doing so makes them less rational by classical standards. At least with respect to the canonical critical thinking capabilities of producing arguments for one's views and producing counterarguments to them, then, interactionism should motivate critical thinking educators to transcend individualism. The best answer to laziness and bias may be other people.

References

- Abrami, P.C., Bernard, R.M., Borokhovski, E., Waddington, D.I., Wade, C.A., and T. Persson. 2015. Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research* 85(2): 275–314.
- Allen, M., Berkowitz, S., Hunt, S., and A. Louden. 1999. A meta-analysis of the impact of forensics and communication education on critical thinking. *Communication Education* 48: 18-30.
- Anderson, R., Chinn, C., Waggoner, M., and K. Nguyen, K. 1998. Intellectually stimulating story discussions. In *Literacy for all*, eds. J. Osborn and E Lehr, 170-196. New York: Guilford.
- Anderson, R., Chinn, C., Chang, J., Waggoner, M., and H. Yi. 1997. On the logical integrity of children's arguments. *Cognition and Instruction* 15: 135-167.
- Anderson, R., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., and Reznitskaya, A., et al. 2001. The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction* 19: 1-46.
- Apple, M. W. 1979. *Ideology and curriculum*. New York: Routledge & Kegan Paul.

- Ariely, D. 2008. *Predictably irrational: The hidden forces that shape our decisions*. New York: HarperCollins.
- Arkes, H. 1981. Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology* 49: 323-330.
- Arkes, H.R. 1991. Costs and benefits of judgement errors: Implications for debiasing. *Psychological Bulletin* 110(3): 486-498.
- Axtell, G. 2017. Thinking twice about virtue and vice: Philosophical situationism and the vicious minds hypothesis. *Logos and Episteme* 8(1): 7-39.
- Beaulac, G. and T. Kenyon. 2018. The scope of debiasing in the classroom. *Topoi* 37(1): 93-102.
- Behar-Horenstein, L.S., and L. Niu. 2011. Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning* 8(2): 23-41.
- Bishop, M. A. and J.D. Trout. 2005. *Epistemology and the psychology of human judgment*. New York: Oxford University Press.
- Boudry, M., Vlerick, M., and R. McKay. 2015. Can evolution get us off the hook? Evaluating the ecological defense of human rationality. *Consciousness and Cognition* 33: 524-535.
- Burbules, N.C. and R. Berk. 1999. Critical thinking and critical pedagogy: Relations, differences, and limits. In *Critical theories in education*, eds. T.S. Pokewitz and L. Fendler. New York: Routledge.
- Chinn, C., and R. Anderson. 1998. The structure of discussions that promote reasoning. *Teachers College Record* 100: 315-368.
- Cowley, M. and R.M.J. Byrne. 2005. When falsification is the only path to truth. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, eds. B.G. Bara, L. Barsalou, and M. Bucciarelli, 512-517. Mahwah, NJ: Lawrence Erlbaum Associates,
- Ennis, R. H. 1995 *Critical Thinking*. Upper Saddle River, NJ: Prentice Hall.
- Freire, P. 1970. *Pedagogy of the oppressed*. New York: Seabury Press.
- Freire, P. 1973. *Education for critical consciousness*. New York: Seabury.

- Todd, P. M. & G. Gigerenzer. 2007. Environments that make us smart. *Current Directions in Psychological Science* 16(3): 167-171.
- Gigerenzer, G. 2000. *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.
- Gigerenzer, G., and P.M. Todd. 1999. Fast and frugal heuristics: The adaptive toolbox. In *Evolution and cognition. Simple heuristics that make us smart*, eds. G. Gigerenzer, P. M. Todd, and The ABC Research Group, 3-34. New York, NY, US: Oxford University Press.
- Goldstein, D. and G. Gigerenzer. 2002. Models of ecological rationality: The recognition heuristic. *Psychological Review* 109(1): 75-90.
- Giroux, H. A. 1988. *Teachers as intellectuals: Toward a critical pedagogy of learning*. South Hadley, MA: Bergin Garvey.
- Hoch, S. J. 1985. Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11: 719-731.
- Johnson, D., Johnson R, and K. Smith. 2014 Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in College Teaching*, 25(3-4): 85-118.
- Kahneman, D. 2011. *Thinking fast and slow*. New York: Farrar, Straus and Giroux.
- Kenyon, T. 2008. *Clear thinking in a blurry world*. Toronto: Thomson Nelson.
- Kenyon, T. and G Beulac. 2014. Critical thinking education and debiasing. *Informal Logic* 34(4): 341-363.
- Koriat, A., Lichtenstein, S., and B. Fischhoff. 1980. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory* 6(2): 107-118.
- Kuhn, D. 1991. *The skills of arguments*. Cambridge, UK: Cambridge University Press.
- Kuhn, D. and W. Udell. 2003. The development of argument skills. *Child Development* 74(5): 1245-1260.
- Larrick R. 2004. Debiasing. In *The Blackwell handbook of judgment and decision making*, eds. D. Koehler and N. Harvey, 316-337. Oxford: Blackwell Publishing.

- Lipman, M. S. 1991. *Thinking in education*. New York: Cambridge University Press.
- Lipman, M. S. 2003. *Thinking in education*. 2nd Ed. New York: Cambridge University Press.
- Laughlin, P. R., Bonner, B. L. and A. G. Miner. 2002. Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes* 88(2): 605-620.
- Laughlin, P. R., Hatch, E. C., Silver, J. S. and L. Boh. 2006. Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *Journal of Personality and Social Psychology* 90(4): 644-651.
- Laughlin, P. R., Zander, M. L., Knieval, E. M. & T. S. Tan. 2003. Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective reasoning. *Journal of Personality and Social Psychology* 85(4): 684-689.
- Lerner, J. S., and P. E. Tetlock. 1994. Accountability and social cognition. In *Encyclopedia of human behavior* Vol. 1, ed. V. S. Ramachandran, 3098-3121. San Diego, CA: Academic Press.
- Lerner, J.S. and P. E. Tetlock. 1999. Accounting for the effects of accountability. *Psychological Bulletin* 125(2): 255-275.
- Lilienfeld, S. O., Ammirati, R., and K. Landfield. 2009. Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science* 4(4): 390-398.
- Lord, C. G., Lepper, M. R., and E. Preston. 1984. Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology* 47(6): 1231-1243.
- Mannes, A. E., Larrick, R. P., and J. B. Soll. 2012. The social psychology of the wisdom of crowds. In *Frontiers of social psychology. Social judgment and decision making*, ed. J. I. Krueger, 227-242. New York: Psychology Press.
- Marshall, D. and B. Weatherson. 2018. Intrinsic vs. extrinsic properties. In *The Stanford Encyclopedia of Philosophy* (Spring 2018 edition), ed. E. N. Zalta. URL accessed 12 August, 2019: <https://plato.stanford.edu/archives/spr2018/entries/intrinsic-extrinsic/>;

- Maynes, J. 2015. Critical thinking and cognitive bias, *Informal Logic* 35(2): 183-203.
- Maynes, J. 2017. Steering into the skid: On the norms of critical thinking, *Informal Logic* 37(2): 115-128.
- Mercier, H. 2017. Confirmation bias—Myside bias. In *Cognitive illusions: Intriguing phenomena in thinking, judgment and memory*, ed. R. F. Pohl, 99–114. New York: Routledge/Taylor & Francis Group.
- Mercier, H., Boudry, M., Paglieri, F. and E. Trouche. 2017. Natural-born arguers: Teaching how to make the best of our reasoning abilities. *Educational Psychologist*, 52(1): 1-16.
- Mercier H., and D. Sperber. 2017. *The enigma of reason*. Cambridge: Harvard University Press.
- Mercier H., and D. Sperber. 2011. Why do humans reason? Arguments for an argumentative theory. *Behavioral & Brain Science* 34(2): 57-74.
- Merritt, M., and S. Varga, S. (Eds). 2013. Special issue: Socially extended cognition. *Cognitive Systems Research*, 25-26: 1-72.
- Morris, J. A. 2017. *Effect of collaborative learning and direct instruction on myside bias* (doctoral dissertation). Urbana-Champaign: University of Illinois at Urbana-Champaign.
- Moshman, D. and M. Geil. 1998. Collaborative reasoning: Evidence for collective rationality. *Thinking and Reasoning* 4(3): 231-248.
- Mussweiler, T. Strack, F. and T. Pfeiffer. 2000. Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin* 26(9): 1142-1150.
- Nussbaum, E. M., and C. S. C. Asterhan. 2016. The psychology of far transfer from classroom argumentation. In *The psychology of argument: Cognitive approaches to argumentation and persuasion*, 407-423. London: College Publications
- Paul, R. W. 1987. Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In *Series of books in psychology. Teaching thinking skills: Theory and practice*, eds. J. B. Baron and R. J. Sternberg, 127-148. New York: W H Freeman/Times Books/ Henry Holt & Co.

- Perkins, D. N. 1985. Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology* 77: 562–571.
- Reznitskaya, A., Anderson, R. C., and L. J. Kuo. 2007. Teaching and learning argumentation. *Elementary School Journal* 107: 450-472.
- Reznitskaya, A., and R. C. Anderson. 2002. The argument schema and learning to reason. In *Comprehension instruction*, eds. C. C. Block and M. Pressley, 319–334. New York: Guilford.
- Reznitskaya, A., Anderson, R., McNurlen, B., NguyenJahiel, K., Archodidou, A., and S. Kim. 2001. Influence of oral discussion on written argument. *Discourse Processes* 32: 155-175.
- Schraw, G. 1998. Promoting general metacognitive awareness. *Instructional Science* 26: 113-125.
- Sharp, A. M. 2014. The other dimension of caring thinking. *Journal of Philosophy in Schools* 1(1): 16-21.
- Sharp, A. M. 1987. What is a 'community of inquiry'? *Journal of Moral Education* 16(1): 37-45.
- Siegel, H. 1988. *Educating reason: Rationality, critical thinking, and education*. New York: Routledge.
- Soll J. B., Milkman K. L., and J. W. Payne. 2015. A user's guide to debiasing. In *Blackwell Handbook of Judgment and Decision Making* (2nd edition), eds. G. Wu and G. Keren. New York: Wiley.
- Sunstein, C. R. and R. Hastie. 2014. *Wiser: Getting beyond group-think to make groups smarter*. Cambridge: Harvard Business Review Press.
- Sunstein, C. R. and R. Hastie. 2008. Four failures of deliberating groups. *John M. Olin Program in Law and Economics Working Paper No. 401*. Chicago: University of Chicago.
- Surowiecki, J. 2004. *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York, NY, US: Doubleday & Co.
- Tetlock, P.T., Skitka, L., and R. Boettger. 1989. Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology* 57(4): 632-40.

- Thagard, P. 2011. Critical thinking and informal logic: Neuropsychological perspectives. *Informal Logic* 31(3): 152-170.
- Thayer-Bacon, B. J. 2000. *Thinking constructively: Transforming critical thinking*. New York: Teachers College Press.
- Tiruneh, D. T., Verburgh, A., and J. Elen. 2014. Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *Higher Education Studies* 4(1): 1-17.
- Tomasello, M. 2009. *The origins of human communication*. Cambridge: MIT Press.
- Stanovich, K. E., West, R. F., and M. E. Toplak. 2013. Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science* 22(4): 259-264.
- Trouche, E., Johansson, P., Hall, L., and H. Mercier. 2016. The selective laziness of reasoning. *Cognitive Science* 40(8): 2122-2136.
- Trouche, E., Sander, E., and H. Mercier. 2014. Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General* 143, 1958-1971.
- Walton, D., Reed, C. and F. Macagno. 2008. *Argumentation schemes*. Cambridge: Cambridge University Press.
- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Thomas W. Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330(6004): 686-688.