The definition of an argument given by M&S is standard: A set of statements, one of which is the conclusion, which is supposed to be epistemically supported by the other statements, called the *premises*. The *content* of an argument refers to the propositions that are expressed by the premises and conclusion, whereas the *structure* of the argument refers to the way the premises work together to support the conclusion. Successfully understanding an argument consists in being able to identify both the content and the structure of the argument: the conclusion, the premises, and the particular way the premises support the conclusion (e.g., whether the premises are linked or convergent). Successfully evaluating an argument consists in being able to assess the content (i.e., determine whether the premises are true) and the structure (i.e., determine whether, assuming that they are true, the premises actually do support the conclusion). Finally, successfully constructing an argument consists in being able to supply true premises and specify how those premises work together to support the conclusion. Although structure and content are both relevant for all three activities, they are relevant in different ways, and so great care is required (but not always taken) in designing experimental tasks that appropriately test them.

Problematic empirical evidence arises for all three: argument understanding, argument evaluation, and argument production. For the first process, there actually seems to be scant research in the area of argument understanding. The little research that does exist in this area is mixed. Some studies (e.g., Ricco 2003, cited by M&S) suggest that for simple arguments, adults can, when prompted, differentiate between linked and convergent arguments. Other studies, however, suggest that, even for simple arguments, untrained college students can identify the conclusion but without prompting are poor at both identifying the premises and how the premises support the conclusion (Harrell 2006; 2008; 2011).

Second, argument evaluation is usually loosely, and only implicitly, defined as being able either to identify reasoning fallacies or to differentiate reasonable arguments from unreasonable ones. The research on argument evaluation seems mixed, at best. In particular, a number of systematic biases have been found. When witnessing an argument from the outside, participants' judgment of the burden of proof depends on who speaks first (Bailenson & Rips 1996, cited by M&S), and participants routinely mistake innocuous repetition for circularity (Rips 2002, cited by M&S). When participating in an argument themselves, participants tend to reason less well than when witnessing an argument (Neuman et al. 2006; Thompson et al. 2005b; both cited by M&S).

Finally, in many of these studies, the perception by the researchers that participants were able to "build complex arguments" (sect. 2.2, para. 3) is vague or ambiguous. Producing an argument is importantly different from, for example, mere fact gathering, but the research focuses almost exclusively on nothing more complex than the listing of reasons to believe. Even for this simple kind of argument production, studies suggest that both low- and high-cognitive-ability participants have difficulty producing evidence for a claim (Sá et al. 2005, cited by M&S).

Contrary to the claims by M&S, a wide literature supports the contention that the particular skills of understanding, evaluating, and producing arguments are generally poor in the population of people who have not had specific training and that specific training is what improves these skills. Some studies, for example, show that students perform significantly better on reasoning tasks only when they have learned to identify premises and conclusions (Shaw 1996, cited by M&S) or have learned some standard argumentation norms (Weinstock et al. 2004, cited by M&S). M&S may be correct that some of these negative results arise because the stakes are too low, but many studies that show improvements from specific training occur in high-stakes environments like a college course (Harrell 2011; Twardy 2004; van Gelder 2005; van Gelder et al. 2004). This suggests that difficulty with understanding, evaluating, and producing arguments may be a deeper feature of our cognition.

The argumentative theory of reasoning applies to scientists and philosophers, too

doi:10.1017/S0140525X10002931

John A. Johnson

Department of Psychology, Pennsylvania State University, DuBois, PA 15801.
j5j@psu.edu
http://www.personal.psu.edu/~j5j/

Abstract: Logical consistency demands that Mercier and Sperber's (M&S's) argumentative theory of reasoning apply to their own reasoning in the target article. Although they hint that their argument applies to professional reasoners such as scientists and philosophers, they do not develop this idea. In this commentary, I discuss the applicability of argumentative theory to science and philosophy, emphasizing the perils of moral reasoning.

Mercier and Sperber (M&S) argue that the primary evolved function of reasoning is persuasive argumentation. If the primary function of any evolved trait – including reasoning ability – is the same for all members of a species, then it follows that professional reasoners (including scientists and philosophers) are primarily in the business of persuasive argumentation. Furthermore, if M&S's dual-process model of reasoning is accurate, professional reasoners initially arrive at their conclusions by intuitive leaps and only later construct logical arguments to convince others of these conclusions. The notion that scientists and philosophers are more concerned with persuading others that something is true than with discovering truth contradicts the image of scientists and philosophers as dispassionate truth-seekers. This response to M&S's target article aims to develop this subversive implication of their argument.

That M&S's argumentative theory applies to their own reasoning is necessary if their theory is to be consistent. To suggest otherwise is to commit what Little (1972) called the *nonreflexive fallacy*. Yet M&S spend virtually the entire article discussing studies of nonscientists and nonphilosophers, with just the briefest mention of how their theory might apply to professional reasoners. One exception is a reference to reviewers of scientific manuscripts who look for flaws in papers to justify rejection when they do not agree with a paper's conclusion. They also remark near the end of their article that even among scientists the ability to control one's own biases is "uncommon" and "almost freakish" (sect. 6, para. 7).

Perhaps the dearth of examples of professional-reasoning-qua-argumentation is due to space limitations. Or, perhaps there is little empirical research on this topic. Or, perhaps other professional reasoners will not find the theory as congenial as M&S suggest in their concluding sentence. After all, it could be somewhat demeaning to see one's professional activity (reasoning) as equivalent to ordinary squabbling over whether my favorite sports team is better than your favorite sports team. Whereas Little (1972) aims to elevate ordinary people to the status of scientists, M&S appear to be challenging the status of scientists and philosophers as elite thinkers. To suggest that "[s]killed arguers, however, are not after the truth but after arguments supporting their views" (see the M&S abstract) is to challenge the idea that scientists and philosophers are motivated in an unbiased way by pure curiosity about what is true.

I believe that we professional reasoners should accept M&S's humbling view of our activities because it is an accurate description of reality. Yes, we are interested in truth, but we relish the thought of convincing others that we have discovered important truths. I must confess that the M&S thesis was immediately congenial to me because it affirms my own long-held beliefs about how professional reasoners such as scientists and moral philosophers go about their work (Johnson et al. 1988). Observations of the actual behavior of scientific researchers indicate that textbook descriptions of science are highly inaccurate. Scientists do not begin with a thorough set of dispassionate observations about

some aspect of the world, followed by formulation of hypotheses that are tested and immediately abandoned if disconfirmed by data. I propose that the following account is more accurate.

Research for professional reasoners begins with an emotional attraction to certain ideas, an attraction Tomkins (1965) called "ideo-affective resonance." This emotional resonance can cause scientists to cling tenaciously to ideas, even in the face of counter-evidence. In some ways, science resembles legal proceedings in which the very best case for guilt or innocence is presented by uncompromising prosecuting and defense attorneys, respectively. Scientists who resonate to different views clash in conferences and in journals. Each seeks to convince others that he or she is correct. M&S review research indicating that when members of groups holding disparate views debate, each arguing for a different view, "truth wins" (sect. 2.3, para. 1). Perhaps truth does win often enough in jury trials and scientific debates, but as we all know, sometimes it does not. M&S might be expressing unwarranted optimism here.

I want to close my commentary with some observations about moral reasoning. Research by Haidt (2001), mentioned by M&S, and by Joshua Greene (2003) strongly supports a dual-process model wherein people instantaneously decide if an act is "good" and therefore something we "ought" to do by taking note of the immediate, reflexive feelings that emerge when thinking about the act. In the second stage of the dual process, they may attempt to defend their feelings in terms of rational argument. Professional philosophers are much better at the reasoning part of the process, but are still guided initially by emotional reflexes. The immediacy and inevitability of certain emotions (e.g., revulsion on contemplating the torture of a child) can lead philosophers and nonphilosophers alike into making pronouncements such as "That we ought to refrain from torturing children is a moral truth."

But only propositions about *what is the case* can be true or false. Moral pronouncements express reflexive feelings about how we ought to behave and are therefore not truth-apt. "Moral truth" is a category mistake. I have a yet-untested two-part hypothesis about why so many people (including moral philosophers) make this apparent category mistake (Johnson 2007). First, human beings are prone to mistakenly assuming that when they feel a strong and immediate emotion, this is a reliable sign of a self-evident truth. Second, although moral systems evolved because they conferred benefits on all participants (compare M&S's observation that persuasive communication must be sufficiently beneficial to both parties, else the capacity for being persuaded would be selected against and go out of existence), the propensity of a person to be responsive to moral "oughts" can be exploited by someone who benefits at that person's expense. Compare, for example, the persuasiveness of "Give me ten percent of your money because I want it" with "That we have a duty to tithe to the church is a venerable moral truth." Scrutiny of any rhetorical effort is wise, particularly those in the moral domain.

True to the power of one? Cognition, argument, and reasoning

doi:10.1017/S0140525X10002992

Drew Michael Khlentzos and Bruce Stevenson

Language and Cognition Research Centre, Psychology, School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale 2351, Australia.

dkhlentz@une.edu.au **bstevens@une.edu.au**
http://www.une.edu.au/staff/dkhlentz.php
http://www.une.edu.au/staff/bstevens.php

Abstract: While impressed by much of what Mercier & Sperber (M&S) offer through their argumentative hypothesis, we question whether the specific competencies entailed in each system are adequate. In particular, whether system 2 might not require independent reasoning capabilities. We explore the adequacy of the explanations offered for confirmation bias and the Wason selection task.

For Mercier and Sperber (M&S), what appears as poor reasoning is actually appropriate argument – social dialogue facilitates reasoning by prompting agents to formulate arguments and defend them from objections. M&S propose a dual-process model with system 1 (S_1) a consortium of inference mechanisms and system 2 (S_2), an S_1 apologist. We identify some features we think require clarification and provide alternative interpretations of phenomena used by M&S to support their model.

If S_1 generates conclusions without revealing their derivation (modular-like), then where does S_2 acquire the competence to support these arguments? What type of reasoning is required for it to construct these arguments, or does it run data back through S_1 for a reasoned result? Related to this is the issue of argumentative contexts which trigger S_2. These appear to be richer in information, creating a potential confound for the argumentative hypothesis: Is it the argumentative feature or the increased information that is critical?

The social psychology findings M&S adduce to support their view present a puzzle for it: How can truth win out amongst sophistical S_2s committed *not* to discovering the facts but to defending S_1's representation of them? Convergence-on-truth suggests there's more to S_2 than defence of S_1. One alternative views S_2 as a dynamic, defeasible reasoner that sifts through S_1 outputs, *independently* generating conclusions to be updated in the light of new information.

Presumably S_1 must support probabilistic as well as deductive inferences. In which case, *some* regulatory role for S_2 is inescapable. Suppose S_1 has both deductive and probabilistic mechanisms and these produce compatible results with input X both deductively entailing and probabilistically supporting Y. Imagine new evidence E emerging that undermines Y so that X + E makes Y *not* probable. Nonetheless, E cannot affect the derivation of Y from X. So X + E still entails Y. Whence S_2 has to decide whether to defend Y since it is derivable from X + E or surrender Y as X + E makes Y improbable. How would it make this decision?

Consider now M&S's views on *confirmation bias*. M&S deny confirmation bias is a flaw in reasoning. Yet if the aim of each agent's S_2 is to persuade others, confirmation bias would just polarize views with no agent prepared to listen to another's arguments. Alternatively, if each S_2 defends an agent's beliefs against objections, amassing evidence for those beliefs is important but anticipating likely objections and preparing a defence is no less so. Relative to aims of persuasion or defence, then, confirmation bias registers as a fault in reasoning.

Compare an M&S-styled S_2-reasoner Aaron with a defeasible S_2-reasoner Belle. Aaron is convinced the river mussels are good to eat since he's eaten them the past five days. Belle felt ill after eating them the day before. She advises Aaron to refrain. Aaron's S_2 considers positive evidence and discounts negative evidence. So Aaron eats the mussels and falls ill. In contrast, Belle's S_2 constructs fast generalizations on the fly. Having eaten them for four days, Belle inferred (G) *the mussels are good to eat*. But now her S_2 enables Belle to adopt a position appropriate to the evolving evidence. The crucial difference between Aaron and Belle is this: Were they to swap roles, Belle would feel no internal pressure from *her* S_2 to eat the mussels (unlike Aaron from his): Evidence someone else fell ill can prompt a defeasible reasoner to update (G) as disconfirming and confirming evidence are weighted equally. Whilst M&S's model allows S_1 to update information, *reasoning* to a new conclusion (belief revision) appears anomalous.

Does the *argumentative hypothesis* yield the best explanation of reasoning performance? Take the Wason selection task. M&S claim that when agents are asked to assess the truth of (W) *If there's a vowel on one side of a card, there's an even number on its other side* for an E, K, 4, 7 array, their S_1 matches cards to