The problems that generate the rationality debate are too easy, given what our economy now demands

Selmer Bringsjord and Yingrui Yang

Department of Philosophy, The Minds and Machines Laboratory, Psychology and Cognitive Science, and Department of Computer Science, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180; **{selmer;yangyri}@rpi.edu**
www.rpi.edu/~brings

Abstract: Stanovich & West (S&W), following all relevant others, define the rationality debate in terms of human performance on certain well-known problems. Unfortunately, these problems are very easy. For that reason, if System 2 cognition is identified with the capacity to solve them, such cognition will not enable humans to meet the cognitive demands of our technological society. Other profound issues arise as well.

The rationality debate revolves around a set of problems, nearly all of which, of course, are well known to the participants in this debate. But all these problems are, to put it bluntly, very easy. This fact – to which the researchers who have hitherto defined the debate are apparently oblivious – has far-reaching consequences, as we begin to explain in this commentary.

To save space, we focus here upon deductive reasoning, and specifically upon syllogistic reasoning. We label a logic problem as "very easy" if there is a simple, easily taught algorithm which, when followed, guarantees a solution to the problem. Normal cognizers who take an appropriate first course in symbolic logic can master this algorithm: Represent a syllogism in accordance with Aristotle's A/E/I/O sentences, cast this representation in first-order logic (FOL), inspect the formalization to see if a proof is possible, carry out the proof if it is, or carry out, in accordance with a certain sub-algorithm, a *dis*proof if it isn't. For 14 years, year in and year out, Bringsjord's students have achieved a more than 95% success rate on post-tests given in his "Introduction to Symbolic Logic" course, in which they are asked to determine whether or not syllogisms are valid. This includes syllogisms of the sort that S&W report subjects to be befuddled by. As an example, consider the "challenging" syllogism S&W present:

(1) All mammals walk.
(2) Whales are mammals.
Therefore: (3) Whales walk.

Each of these sentences is an A-sentence (All A are B):

(1′) All M are A.
(2′) All W are M.
Therefore: (3′) All W are A.

So in FOL we have:

(1″) $\forall x\ (Mx \rightarrow Ax)$
(read: for all x, if x is an M, then x is an A)
(2″) $\forall x\ (Wx \rightarrow Mx)$
Therefore: (3″) $\forall x\ (Wx \rightarrow Ax)$

The proof now runs as follows: Let *a* be an arbitrary thing. We can instantiate the quantifiers in (1″) and (2″) to infer $Ma \rightarrow Aa$ and Wa (Ma), respectively. We can then use hypothetical syllogism (a "chain rule") to conclude $Wa \rightarrow Aa$. Since *a* was arbitrary, from this we can conclude by universal introduction $\forall x\ (Wx \rightarrow Ax)$. QED.

For every formally valid syllogism, the corresponding proof can be generated by such simple mechanical means. What about formally invalid syllogisms? Producing disproofs is here once again a matter of following a trivial algorithm. To show this, consider an example from Johnson-Laird & Savary (1995). When asked what can be (correctly) inferred from the two propositions

(4) All the Frenchmen in the room are wine-drinkers.
(5) Some of the wine-drinkers in the room are gourmets.

most subjects respond with

Therefore: (6) Some of the Frenchmen in the room are gourmets.

Alas, (6) cannot be derived from (4) and (5), as can be seen by inspection after the problem is decontextualized into FOL, and chaining is sought.

But Bringsjord's students, trained to use both the algorithm above, and therefore the sub-algorithm within it for generating disproofs, and nothing else, not only cannot make the erroneous inference, but *can also* prove that the inference is erroneous. Here's why. The Aristotelean form consists of one A-sentence and two E- sentences (Some A are B):

(4′) All F are W.
(5′) Some W are G.
Therefore: (6′) Some F are G.

In FOL this becomes

(4″) $\forall x\ (Fx \rightarrow Wx)$
(5″) $\exists x\ (Wx\ \&\ Gx)$
Therefore: (6″) $\exists x\ (Fx\ \&\ Gx)$

Notice, first, that neither Wa nor Ga can be used to chain through $Fa \rightarrow Wa$ to obtain the needed Fa. Next, for a disproof, imagine worlds whose only inhabitants can be simple geometric shapes of three kinds: dodecahedrons (dodecs), cubes, and tetrahedrons (tets). Suppose now that we fix a world populated by two happy, small dodecs, two happy, large cubes, and two medium tets. In this world, all dodecs are happy (satisfying premise [4″]), there exists at least one happy, large thing (satisfying premise [5″]), and yet it is not the case that there is a large dodec (falsifying proposition [6″]). Students in Bringsjord's logic course, and in logic courses across the world, mechanically produce these disproofs, often by using two software systems that allow for such worlds to be systematically created with point-and-click ease. (The systems are Hyperproof and Tarski's World, both due to Barwise & Etchemendy 1984; 1999.) One of us has elsewhere argued that the appropriate pedagogical deployment of these two remarkable systems substantiates in no small part the neo-Piagetian claim that normal, suitably educated cognizers are masters of more than System 2 cognition at the level of FOL (Bringsjord et al. 1998). Whether or not Bringsjord is right, it's hard to see how S&W consider the neo-Piagetian response to the normative/descriptive gap. They consider a quartet of proposed explanations – fundamental irrationality, performance errors, computational limitations, misconstrual of problem. But why can't the gap be explained by the fact that most people are just uneducated? (In his first-round commentary, Zizzo [2000] mentions the possibility of teaching logic on a mass scale, but then seems to reject the idea. Actually, by our lights, that's exactly what needs to be done in order to meet the demands of our high-tech economy.)

Now we know that S&W, in responding to Schneider's (2000) first-round commentary, point out that the correlation between heuristics and biases tasks and training in mathematics and statistics is negligible (Stanovich & West 2000, p. 705). But this is irrelevant, for two reasons. First, S&W ignore Schneider's specific claim about syllogisms, and (tendentiously?) zero in on her claim that suitable education can cultivate a cognition that leads to higher SAT scores. What Schneider says about syllogisms is that some people can effortlessly and accurately assess them (albeit via System 1 cognition in her cited cases). Second, the issue, in general, is whether *specific* training has an effect on performance. Few math courses (traditionally, none before analysis) at the undergraduate (and even, in more applied departments, at the graduate) level explicitly teach formal deductive reasoning, and many first logic courses are merely courses in *informal* reasoning and so-called critical thinking – courses, therefore, that don't aim to teach decontextualization into some logical system. This is probably why the problem of moving from mere problem solving in mathematics to formal deductive reasoning (a problem known as "transition to proof"; Moore 1994) plagues nearly all students of math, however high their standardized test scores; and why, in general, there is little correlation between math education and the solving of those problems in the rationality debate calling for deductive reasoning. The meaningful correlation would be between subjects who have had two or more courses in symbolic logic and high performance, for example, on (very easy) deductive reasoning problems seen in the rationality debate. We predict that this correlation will be strikingly high. (See also the prediction made by Jou [2000, p. 680] in the first round of commentary, concerning scores on the logical reasoning section of the GRE and normative performance. In this connection, it is probably noteworthy that those who write on logical reasoning in "high stakes" standardized tests invariably have training in symbolic logic.)

We heartily agree with S&W that today's workforce demands rigorous, deoncontextualized thinking on the part of those who would prosper in it. In their response to the first round of commentaries, the authors provide a nice list of relevant challenges (p. 714); let's take just one: deciding how to apportion retirement savings. In our cases, which are doubtless representative, we can choose to set up our 403(b)'s with one of three companies, each of which offers, on the mutual fund front alone, one hundred or so options. One tiny decision made by one fund manager makes syllogistic reasoning look ridiculously simple by comparison, as any of the proofs driving financial knowledge-based expert systems make plain. To assess the future performance of many such managers making thousands of decisions on the basis of tens of thousands of data points, and at least hundreds of declarative principles (*and,* for that matter, an array of rules of inference as well), is not, we daresay, very easy. Logicians can crack syllogisms in seconds, yes. But if you tried to configure your 403(b) in a thoroughly rigorous, decontextualized way, how long did it take you?

Other, arguably even deeper, problems spring from the simplicity of the problems that currently anchor the rationality debate. It seems bizarre to define general intelligence as the capacity to solve very easy problems. For example, Raven's Progressive Matrices, that vaunted "culture-free" gauge of *g*, can be mechanically solved (Carpenter et al. 1990). Once one assimilates and deploys the algorithm, does one suddenly become super-intelligent? Would a computer program able to run the algorithm and thereby instantly solve the problems, be counted genuinely intelligent? Hardly. (For more on this issue, see Bringsjord 2000. And recall Sternberg's continuous complaint that "being smart" in the ordinary sense has precious little to do with solving small, tightly defined test problems, a complaint communicated to some degree in his first-round commentary; cf. Sternberg 2000.)

Another problem arising from the fact that the rationality debate is tied to very easy problems is that psychology of reasoning is thereby structurally unable to articulate theories of robust human reasoning. Mental logic (championed, for example, by Rips 1994) cannot account for disproofs of the sort we gave above (because such disproofs are necessarily meta-proofs carried out outside a fixed set of inference schemas); and mental models theory (Johnson-Laird 1983), which rejects elaborate sequences of purely syntactic inferences, would seem to at least have a difficult time accounting for solutions to the problem we leave you with below (about which we've just given you a hint). What is needed is a theory of human reasoning that partakes of both the proof theoretic and semantic sides of symbolic logic, and the formal metatheory that bridges these two sides. (For a synoptic presentation of all this terrain, in connection to cognition and reasoning, see Bringsjord & Ferrucci 1998. For a theory of human reasoning designed to cover all of this terrain, Mental MetaLogic, see (Yang & Bringsjord, under review.)

Finally, what would be an example of a reasoning problem that *isn't* very easy, and the solving of which might justify confidence that the solver is both poised for success in the high-tech twenty-first century, and genuinely intelligent? Well, here's one; we refer to it as "The Bird Problem": Is the following statement true or false? Prove that you are correct.

(7) There exists something which is such that, if it's a bird, then everything is a bird.