BEHAVIORAL AND BRAIN SCIENCES (2003) **26,** 527–534
Printed in the United States of America

Continuing Commentary

Commentary on **Stanovich, K. E. and West, R. F. "Individual differences in reasoning: Implications for the rationality debate?" BBS 23(5):645–726 [October 2000]**

Abstract of the original article: Much research in the last two decades has demonstrated that human responses deviate from the performance deemed normative according to various models of decision making and rational judgment (e.g., the basic axioms of utility theory). This gap between the normative and the descriptive can be interpreted as indicating systematic irrationalities in human cognition. However, four alternative interpretations preserve the assumption that human behavior and cognition is largely rational. These posit that the gap is due to (1) performance errors, (2) computational limitations, (3) the wrong norm being applied by the experimenter, and (4) a different construal of the task by the subject. In the debates about the viability of these alternative explanations, attention has been focused too narrowly on the modal response. In a series of experiments involving most of the classic tasks in the heuristics and biases literature, we have examined the implications of individual differences in performance for each of the four explanations of the normative/descriptive gap. Performance errors are a minor factor in the gap; computational limitations underlie nonnormative responding on several tasks, particularly those that involve some type of cognitive decontextualization. Unexpected patterns of covariance can suggest when the wrong norm is being applied to a task or when an alternative construal of the task should be considered appropriate.

Freud's dual process theory and the place of the a-rational

Linda A. W. Brakel and Howard Shevrin
Department of Psychiatry, University of Michigan Medical Center, Ann Arbor, MI 48109-0722

Abstract: In this commentary on Stanovich & West (S&W) we call attention to two points: (1) Freud's original dual process theory, which antedates others by some seventy-five years, deserves inclusion in any consideration of dual process theories. His concepts of primary and secondary processes (Systems 1 and 2, respectively) anticipate significant aspects of current dual process theories and provide an explanation for many of their characteristics. (2) System 1 is neither rational nor irrational, but instead a-rational. Nevertheless, both the a-rational System 1 and the rational System 2 can each have different roles in enhancing evolutionary fitness. Lastly, System 1 operations are incorrectly deemed "rational" whenever they increase evolutionary fitness.

We would like to raise two matters concerning the target article by Stanovich & West (2001) (henceforth S&W). The first concerns a dual processing system conceptualized by Freud (1895; 1900; 1911; 1915; 1940/1964). Freud termed his System 1 "the primary process" and his System 2 "the secondary process." Given that Freud's two systems theory predates any cited in the target article by more than 75 years, his theory ought not be omitted. In fact, Freud deserves priority because he anticipated, on the basis of clinical experience and theoretical considerations, the more recent experimentally established dual process findings. The primary process operates in the unconscious, appears early in development, and obeys laws different from the secondary process. These different laws are marked by "a compulsion to associate" (Freud 1895/1964, p. 338). Moreover, as Anna Freud later elaborated:

> when the "primary process" prevails, there is no synthesis of ideas, affects are liable to displacement, opposites are not mutually exclusive and may even coincide, and condensation occurs as a matter of course. (A. Freud 1936, p. 7)

Further, Freud (1915/1964) held that primary process mentation is incapable of representing negation, degrees of certainty, and past or future – primary process is always in the present tense (pp. 186–87). On the other hand, secondary process mentation is capable of representing past and future as well as the present, and constitutes an instrument for testing reality based on standard logic and considerations of truth and falsity.

Thus, in terms of S&W's Table 3 (*BBS* 23(5), p. 659), Freud's primary process, like other examples of System 1 theories, operates as an associative system. The primary process does frequently seem to function in a holistic fashion, although treating the part as a whole also occurs. Like other System 1 theories, the primary process can be quick and inflexible; however, since it is not bound by the constraints of logic and rationality, primary process mentation can also be characterized as highly (even overly) variable. The primary process is operative in implicit and unconscious cognition and it is highly contextualized and personal. Primary process mentation operates on the basis of isolated and often nonessential features rather than on complex relationships among features (Brakel et al. 2000; Brakel 2001). From the standpoint of normative expectations such thinking will seem irrational, either because so much appears to be ignored, or so little is over-interpreted, occasioning a response that seems excessive. The secondary process is quite different in that it is rule-based, decontextualized, and depersonalized. In secondary process mediated judgments and categorization tasks, nonessential features are overlooked in favor of salient relationships among features. The secondary process functions in accord with the basic tenets of logic and as such is vital for explicit learning and inference.

Two other aspects of Freud's conception of the primary process help explain why System 1 is personalized, conversational, and socialized. In Freud's conception, primary process (System 1) is linked closely to basic motivations that have other people as their main objects and serve important personal and interpersonal needs. Although not always the most adaptive (in any sense), the primary process provides the quickest avenue for immediate satisfaction of these personal and interpersonal needs. The second

 0140-525X/03 $12.50

important aspect concerns the developmental dimension. The primary process is called primary because it precedes the secondary process in development. According to some of our own evidence, the primary process appears to prevail until about seven years of age (Brakel et al. 2002). However, even though primary process is then largely replaced by secondary process, the primary process continues to operate unconsciously and may appear consciously under circumstances of stress and conflict. (Brakel & Shevrin, submitted a; submitted b) From a Freudian standpoint, the primary and secondary process systems are co-existent throughout adult life.

The second matter concerns whether the primary process, and the other System 1 operations, need to be considered "rational" in any way – including "evolutionarily rational." We would like to suggest that, (1) System 1 operations, like the primary process, are neither rational nor irrational, but instead *a*-rational; (2) nonetheless, both the a-rational System 1 and the rational System 2 can each have different roles in enhancing evolutionary fitness; (3) System 1 (primary process type) operations are deemed "rational" (incorrectly, in our view) precisely when their operations effect increased evolutionary fitness.

(1) The a-rationality of System 1 can be seen in a certain category mistake made by birds, frogs, and people. A bird responds to a big black cloth just as it does to a bird bigger than itself. A frog swallows BB pellets with the same alacrity as it does bugs and flies. A spider-phobic person responds with dread to a plastic or rubber replica of a spider much as he/she does to a real spider. In all of these instances we are not dealing merely with ambiguous or unclear perceptions. In each instance, the black cloth, BB pellet, and spider replica, even when seen up close and (certainly in the human case and likely in the others) accurately perceived, still occasion the fearful flight, swallowing, and dread reactions. We are dealing with "funny System 1 categories" – something to flee from, something to swallow, something to fear – that include items resembling the truly biologically relevant exemplars only superficially. Individual features, often nonessential ones, determine category membership. Now, while it could be argued that in the human case the individual is aware of the seeming irrationality of his/her response, nevertheless the fearful response to the replica spider persists based on the a-rational category match. Unlike people, birds are not rational and therefore cannot be irrational. Birds make no rational appraisals of the a-rational category matches underlying their urgent flights from big black cloths. Where there is no capacity to choose in a rational fashion, behavior predicated on a-rational principles is not irrational.

(2) Both System 1 and System 2 can play different roles in evolutionary fitness. There is general agreement that employing the basic rules of logic, especially in the service of reality testing, can be useful to individuals who are capable of such rational cognitive feats. Indeed, as pointed out in the target article, such capacities certainly can further individual goals. But clearly, insofar as System 2 operations are invaluable, not just in the psychology lab but in obtaining food, shelter, and mates, having these capacities will also enhance the reproductive fitness of individuals. With respect to System 1 operations, these too can enhance evolutionary success – but precisely because System 1 is not a rational system and therefore does not have the constraints of rationality.

Suppose bugs, whose life cycles are much faster than those of frogs, evolve forms that weigh as much as metal BBs and even taste like them. Frogs behaving on the basis of their System 1 category will eat better and reproduce more successfully than would frogs who could make System 2 inferences like, "If round and weighty and metallic tasting, then not food, don't swallow." Likewise, the bird with the "funny category" titled "flee, it is big!" will be able to avoid more novel non-avian-but-nonetheless-dangerous-things. Further, unless useless flights are too much of an energy drain, the birds with a-rational categories will do better reproductively than "rational" birds, who would, for example, fail to avoid dangerous human-made items such as pieces of steel hoisted skyward in a construction area. Even if useless flights compromise reproductive potential, they are still less damaging than sudden death. In the human case, to the extent that spiders are truly dangerous, it is best to avoid anything that might possibly be a spider – even if the judgment is made under the quick and dirty System 1 – since spider bites can be fatal, and avoiding rubber spiders merely silly.

The phenomenon of stranger anxiety provides a more general example of a possible adaptational advantage of a-rational mentation in people. Although stranger anxiety does not occur in all healthy babies, it does occur in many at around eight months. Infants younger than eight months, largely System 1 thinkers, can be comfortably cared for by any nurturing adult. Not so when certain babies become more capable of even rudimentary System 2 categorizing and infer, "since these are not my parents, I should be wary." (For further examination of the evolutionary role of a-rational mentation, see Brakel 2002).

Note here that both of our examples concerning people demonstrate that evolutionary explanations can provide only necessary, but not sufficient conditions for a-rational mediated behaviors in human beings. On the basis of temperament, some normal babies may have a very mellow version of wariness. And with regard to spider phobias, indeed spiders, rather than butterflies, may have been selected by evolution as a target for phobias; but for each spider phobic there must also be specific psychological experiences that result in the development of a phobia.

(3) When a-rational mechanisms yield success, it is not on the basis of rationality, but because they function with fewer constraints and more category inclusiveness. While we are not claiming that these System 1 operations work most of the time – many pointless flights, nutrition-less swallowings take place – they work just often enough to enhance reproductive fitness and evolutionary success. But when this outcome is thought to demonstrate "evolutionary rationality," more than just a misnomer is involved. It is precisely the "evolutionary a-rationality" that has contributed to enhanced reproductive fitness. In other words, System 1 operations, whether or not successful, do not result from irrationality or error in an otherwise rational system. Rather, the evolutionary utility of System 1 is due to precisely its nonrational or arational principles – principles operating quite differently from System 2 rationality.

As we stated above, Freud posited that although there is developmental progression whereby primary process is largely supplanted by secondary process, primary process continues to operate, often unconsciously, often influencing behavior significantly. Indeed, for Freud, primary process and secondary process systems are co-existent throughout adult life. Now more than a century later, as it becomes clearer that a-rational System 1 and rational System 2 are each important for enhancing evolutionary fitness in different ways, Freud's early conceptualizations of dual processing systems look better and better.

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.

Individual differences transcend the rationality debate

Elizabeth J. Newton[a] and Maxwell J. Roberts[b]

[a]*Department of Human Communication Science, University College London, London, WC1N 1PF, United Kingdom;* [b]*Department of Psychology, University of Essex, Colchester, Essex, CO4 3SQ, United Kingdom.*
liz.newton@ucl.ac.uk mjr@essex.ac.uk

Abstract: Individual differences are indeed an important aid to our understanding of human cognition, but the importance of the rationality debate is open to question. An understanding of the process involved, and how and why differences occur, is fundamental to our understanding of human reasoning and decision making.

The main thesis of Stanovich & West (S&W) is that differences in individuals' performance can be used to cast light on the rationality debate. Even if we accept that this issue is important, and that humans occasionally behave irrationally, we still need clear criteria to identify such behaviour. Responses by themselves are often taken to be sufficient, but these are only informative if the cognitive processes underlying them are also understood. Otherwise, there is little to gain by addressing the rationality question. This problem applies equally to the interpretation of psychometric test scores. Intelligence is a poorly understood construct, and the suggestion that it reflects only working memory capacity is by no means fully accepted.

Although highly intelligent people may be more likely to give normative responses in reasoning and decision-making tasks than less intelligent people, correlations between test score and reasoning performance can occur for a variety of reasons. Hence, differential correlations are not necessarily informative, and the focus on individual differences in terms of outputs rather than processes means that important qualitative differences are overlooked. The most straightforward reason for a correlation between intelligence test score and performance at a reasoning task is that highly intelligent people use the same processes as less intelligent people, but execute them more effectively. However, this merely leads us back to a "cognitive limitations" account of irrational behavior. Alternatively, perhaps highly intelligent people are better able to use *different,* more complex processes. A further possibility is that they are more likely to use *different* processes, but these are simpler, and hence more efficient. Either possibility leads into a debate about whether the strategy selections, rather than the responses themselves, are rational. A resolution depends crucially upon the ability to identify reasoning strategies accurately at the level of the individual. However, even where this is possible, we have to be certain that a suboptimal strategy, that is, one that is linked to poor performance, is really failing because of fundamental flaws. If, instead, a strategy is potentially normative, but too demanding to be executed accurately, then this turns the rationality issue into a debate concerning whether a person has made a strategy choice commensurate with his or her own ability to execute it accurately – and, ultimately, we are again returned to a cognitive limitations explanation of irrational behavior.

Given that people differ in the strategies they use, the rationality debate forces a dichotomy on us: are these choices normative or non-normative? Suppose you are presented with a series of trials, each consisting of compass point directions given together (e.g., one step north, one step east, one step north, one step west, one step south, one step south, one step west, one step north). The task is to determine the end point, relative to the start, after taking the steps. The natural strategy for this task is spatial: The full path is traced in the mind or by using a finger. For the cancellation strategy – a *task specific short-cut* – opposite steps are cancelled, with the remainder forming the correct response. Both strategies are normative: where applied accurately, they will yield the correct response. However, the spatial strategy is slower, less accurate and more demanding to execute. People are often painfully aware of the need to find an alternative. Surprisingly, even amongst university students, cancellation is used only by the minority. This is because it is only available to people with sufficiently high spatial ability to be able to identify the redundant processes of the spatial strategy and delete them, leading to the discovery of cancellation. Hence, people with high spatial ability outperform the rest, not because they are executing the spatial strategy more efficiently, nor because they are better able to use an enhanced spatial strategy in which additional processes increase accuracy, but because they have dispensed with spatial representations altogether, increasing accuracy and minimising effort (see Newton & Roberts 2000; Roberts et al. 1997).

So, are the spatial strategy users irrational for this task? Granted, if we focus on outputs only, they are less accurate than cancellation users; but, as suggested earlier, the rationality debate is only served crudely in this way. Errors when using the spatial strategy are due to capacity limitations in any case. The rationality debate is not served at all by considering whether selected strategies are normative – both will yield correct responses if executed accurately. Are the spatial strategy users less rational because they made an inappropriate choice? No, there is no alternative available to them, and hence no choice. If the spatial strategy users must be branded as irrational, this can only be because they lack the necessary resources to discover the more efficient method. This conceptualisation of irrationality as a lack of creativity may be unappealing to some, but for this example the categorisation of these people as irrational is artificial and driven only by the perceived need to address the rationality question.

Overall, S&W make many valid points. We agree that individual differences are an important aspect of human cognition. But to use them merely to resolve the rationality debate is problematic and neglects their full potential. Issues of how people reason, and how these processes change and develop with experience can be better answered by not being side-tracked in this way.

Authors' Response

The rationality debate as a progressive research program

Keith E. Stanovich[a] and Richard F. West[b]

[a]*Department of Human Development and Applied Psychology, University of Toronto, Toronto, Ontario, M5S 1V6, Canada;* [b]*School of Psychology, James Madison University, Harrisonburg, VA 22807.* **kstanovich@oise.utoronto.ca westrf@imu.edu falcom.jum.edu//~westrf**

Abstract: We did not, as Brakel & Shevrin imply, intend to classify either System 1 or System 2 as rational or irrational. Instrumental rationality is assessed at the organismic level, not at the subpersonal level. Thus, neither System 1 nor System 2 are themselves inherently rational or irrational. Also, that genetic fitness and instrumental rationality are not to be equated was a major theme in our target article. We disagree with Bringsjord & Yang's point that the tasks used in the heuristics and biases literature are easy. Bringsjord & Yang too readily conflate the *ability* to utilize a principle of rational choice with the *disposition* to do so. Thus, they undervalue tasks in the cognitive science literature that compellingly reveal difficulties with the latter. We agree with Newton & Roberts that models at the algorithmic level of analysis are crucial, but we disagree with their implication that attention to issues of rationality at the intentional level of analysis impedes work at the algorithmic level of analysis.

We found much that is congenial to our way of thinking in these three commentaries. For example, we welcome **Brakel & Shevrin**'s points about Freud and dual-process theorizing. Such theorizing of course predates Freud as well, going back at least to Plato. As Plato writes in *The Republic,*

> we may call that part of the soul whereby it reflects, rational; and the other, with which it feels hunger and thirst and is distracted by sexual passion and all the other desires, we will call irrational appetite, associated with pleasure in the replenishment of certain wants. (Cornford 1945, p. 137)

While we welcome Brakel & Shevrin's addendum, we take it as understood that the purpose of our paper was not the historical exegesis of dual-process notions. Some historian really does need to do a treatise tracing dual process ideas from Plato, through Freud, to the cognitive revolution (e.g., Evans & Wason 1976; Shiffrin & Schneider 1977), but this was not our purpose. Our argument depended only upon common assumptions of these theories and not on nuanced differences or historical relationships.

There is much in **Brakel & Shevrin**'s characterization of System 1 and System 2 that we agree with. For example, we agree (as do most of the dual-process theorists that we cite in the target article) that System 1 processing is not supplanted by System 2 processing with development, but rather, that both types of processing continue to operate in parallel. However, there are some points of misinterpretation as well. Brakel & Shevrin seem to imply that we are labelling *systems* as rational or irrational, but this is not the case. Instrumental rationality (what was termed normative rationality in our target article) is assessed at the organismic level, not at the subpersonal level. Neither System 1 nor System 2 are themselves inherently rational or irrational.

We focused in the target article and elsewhere (e.g., Stanovich & West 2003) on situations where System 1 functioning served to disrupt the pursuit of instrumental rationality (if not overridden by System 2 processes). But we were also clear to note in the target article that "It must be stressed though that in the vast majority of mundane situations, the evolutionary rationality embodied in System 1 processes will *also* serve the goals of normative rationality" (Stanovich & West 2000, p. 661); and in our Authors' Response we repeated that "we made it clear in the target article that in most cases the goals of Systems 1 and 2 will coincide and that System 1 processes will often *also* serve the goal of normative rationality" (p. 708). So, System 1 serves the organism most of the time by facilitating instrumental rationality, but sometimes disrupts the pursuit of instrumental rationality and must be overridden by System 2.

Thus, System 1 is not appropriately characterized *itself* as being either inherently rational or irrational – a point we feel we made clear in the original target article. Furthermore, the same is true of System 2. It can instantiate rules of rational thought which facilitate maximal goal satisfaction (our emphasis in the target article), but it can also instantiate ideas and rules (memes, in the view of Dennett 1991 and Blackmore 1999) that impede the organism's pursuit of instrumental rationality – a theme we did not emphasize in the target article, but have stressed in subsequent publications (Stanovich 2004; Stanovich & West 2003). Thus, System 2 likewise should not be characterized as inherently rational or irrational, since it too is a subpersonal entity.

Brakel & Shevrin seem to have been confused by our use of the term evolutionary rationality, but here the fault might be ours. Our use of the term in the target article was perhaps too clever by half. The term was coined as an indirect tweak at the evolutionary psychologists who conflate behavior serving genetic fitness with behavior that is instrumentally rational (a major theme in our book-length treatments of these issues; Stanovich 1999; 2004). The terms evolutionary rationality (behavior serving genetic fitness) and normative rationality (instrumental rationality) were meant to separate these two. For example, Over (2000), in his critique of work on fast and frugal heuristics (e.g., Todd & Gigerenzer 2000), makes use of our distinction in exactly the way we intended. Nevertheless, we acknowledge that the term evolutionary rationality may have invited people to conflate just the distinction that we wished to emphasize (as Brakel & Shevrin seem to have done). Thus, in a new book by one of us (Stanovich 2004) – which is largely devoted to working out the implications of mismatches between behavior serving the interests of replicators in the environment of evolutionary adaptation and current instrumental rationality for the organism – the term is omitted in favor of stating exactly what it is, fitness at the level of the gene.

Brakel & Shevrin's characterization of rationality seems to be overly tied to a conception that emphasizes conscious reasoning according to logical rules. Our conception of rational choice and thought is informed by the much more general conception of rationality in modern cognitive science and decision theory. This difference is apparent in Brakel & Shevrin's statement that birds cannot be rational or irrational. Decision scientists disagree. Much work has been done on whether animals (some as simple as bees) satisfy the strictures of axiomatic utility theory (see Kagel 1987; Real 1991; Shafir et al. 2002). Most cognitive scientists would agree with Millar (2001) that higher-order representation is necessary for something to be a rational *agent,* but it is not necessary for something to be called a rational animal.

As Meliorists, we share **Bringsjord & Yang**'s concern for emphasizing the effects of education on reasoning. We disagree, however, that the problems studied by researchers in the rationality debate are easy, and by implication trivial. We do not share their definition of what is an easy reasoning problem (in Bringsjord & Yang's view, a problem is easy "if there is a simple, easily taught algorithm which, when followed, guarantees a solution to the problem"). Many principles of rational thought can be acquired, but without the dispositions and/or skills necessary to appreciate the applicability of the principles. Many statistics instructors experience frustration with students who learn principles such as the law of large numbers or regression to the mean but cannot think to apply these principles in situations where they are applicable. This is why our own research group and many others (e.g., Klaczynski et al. 1997; Newstead et al. 2002; Perkins 1995; Sa et al. 1999; Schommer 1990;Sinatra & Pintrich 2003; Sternberg 1997) have focused not only on the principles themselves, but also on the cognitive dispositions that facilitate their actual *use* in real contexts. The common distinction in the critical thinking literature between abilities and dispositions is important.

Thus, we do not agree that, just because the principle *behind* a task in the heuristics and biases literature is easily taught, the problem itself is easy, and that it is not relevant to functioning in the modern world. Many of the axioms of rational choice (e.g., transitivity, the sure-thing principle, independence of irrelevant alternatives) are quite easy to apply and teach, but the decision theory literature is littered with dozens of studies showing that the ability to appropriately apply the (admittedly very simple) principles can be a difficult, though important, skill to acquire. As Shafir et al. (1993) note,

> it has been repeatedly observed that the axioms of rational choice which are often violated in non-transparent situations are generally satisfied when their application is transparent. . . . These results suggest that the axioms of rational choice act as compelling arguments, or reasons, for making a particular decision when their applicability has been detected, not as universal laws that constrain people's choice. (p. 34)

Difficulty in seeing the applicability of very simple choice axioms in real-life tasks has been amply demonstrated in the decision theory literature. This is why the rational thinking skills involved should in no way be characterized as simple or trivial (even *though* education in the skills can improve them). Because of the failure to apply some basic choice axioms, people choose less effective medical treatments; people fail to accurately assess risks in their environment; information is misused in legal proceedings; parents fail to vaccinate their children; billions of dollars are wasted on quack medical remedies; and costly financial misjudgments are made (e.g., Baron 1998; Bazerman et al. 2001; Belsky & Gilovich 1999; Dawes 2001; Kahneman & Tversky 2000; Margolis 1996; Russo & Schoemaker 2002).

As we indicated in the Authors' Response to other commentators who raised the issue of the importance of a process analysis of the tasks used in cognitive science (Stanovich & West 2000), we agree with **Newton & Roberts** that a fully explicated model at the algorithmic level of analysis is a crucial part of most cognitive science endeavors. We reiterate that we have worked at just such a level of analysis in another task domain of cognitive psychology for over two decades (Stanovich 2000; West & Stanovich 1978; 1986). We disagree, however, with the implication (in phrases like "there is little to gain by addressing the rationality question") in the Newton & Roberts commentary that it is a zero-sum game – that a focus at the intentional level of analysis precludes work at the algorithmic level. Our Authors' Response pointed to the venerable tradition in cognitive science (Anderson 1990; 1991; Levelt 1995; Marr 1982; Newell 1982; Oaksford & Chater 1995) which supports the notion that there can be synergistic interplay between levels. Indeed, one could view the interdisciplinary field of cognitive science as reflecting an attempt to integrate sciences focused on the algorithmic level of analysis (e.g., psychology) with sciences focused on the intentional level (e.g., anthropology, economics). Thus, although we wholeheartedly agree that individual difference analyses at the algorithmic level of analysis – of the type that Newton & Roberts are conducting in their ongoing research program – are of immense importance, we disagree with their denigration of individual difference analyses at the intentional level.

In fact, the generic dual-process models that we discuss in the target article represent the beginnings of an algorithmic understanding of the source of irrational responding. Other investigators have been refining the specifics of this generic process explanation (e.g., Evans 2002; Sloman 2002; Sloman & Rips 1998; Slovic et al. 2002) and some neurophysiological work on it has also appeared (Goel & Dolan 2003). Kahneman and Frederick (2002) describe dual-process explanations of many effects in the heuristics and biases literature that never would have become objects of attention except for the intentional-level focus on the goal-thwarting properties of the typical response on the task – that is, never would have been objects of attention *except* for the rationality debate.

We also disagree with how they frame their discussion of the compass point task. For reasons that are related to our earlier remarks on the **Brakel & Shevrin** commentary, we do not think the question of whether a certain (internal) strategy is rational or irrational is well formed. We do not believe the term rationality applies to subpersonal entities. Rationality concerns the actions of an entity in its environment that serve its goals. One could, of course, extrapolate the notion of environment to include the interior of the brain itself, and then talk of a submodule that chose strategies rationally or not. This move creates two problems. First, what are the goals of this subpersonal entity – what are its interests that its rationality is trying to serve? This is unclear in the case of a subpersonal entity. Second, such a move regresses all the way down. We would need to talk of a neuron firing being either rational or irrational ("turtles all the way down!"). It was a version of this mistake that we invited by our use of the term evolutionary rationality. It was, of course, not means-ends rationality we had in mind for a gene, but the optimization of its fitness in a biological sense.

The task in question is not a good example of any of the points relevant to our target article. Unlike many tasks in the heuristics and biases literature, the normative response in the compass task is not in dispute. A correlational analysis of the type we applied to the former would reveal a fairly mundane result. Fewer errors on the compass task would be made by individuals utilizing the cancelling strategy and, as **Newton & Roberts** note, they would be subjects of higher ability. This would yield a correlation between ability and the normative response – a correlation utterly expected on tasks for which there is no dispute about the normative response. Asking whether the spatial strategy is rational or not is a category mistake. The spatial strategy is less efficient, and thus subpar performance on the task due to the use of the spatial strategy represents a computational limitation in our taxonomy, albeit of a somewhat different type than that discussed in our target article. However, on page 239 of our book-length treatment (Stanovich 1999) we discuss different types of computational limitations that would encompass instances more similar to that occurring in this example.

The research program sketched by **Newton & Roberts** seems indeed an extremely useful one, but likewise, our use of the rationality debate to discuss a mix of individual differences at the algorithmic and intentional level has born fruit in the study of belief bias (Stanovich & West 1997), schizophrenia (Oaksford & Sellen 2000), disjunctive reasoning (Toplak & Stanovich 2002), developmental trends (Klaczynski 2001; Kokis et al. 2002), conceptual change (Southerland & Sinatra 2003), and discontuities between intelligence and rational behavior (Sternberg 2002). Many other researchers (e.g., Elio 2002; Evans & Over 1996; Kuhberger 2002; Over 2002), like us, see the rationality debate as a progressive research program.

References

[Note: The letter "r" appearing before authors' initials refers to response article references]

Anderson, J. R. (1990) *The adaptive character of thought.* Erlbaum. [rKES]
Anderson, J. R. (1991) Is human cognition adaptive? *Behavioral and Brain Sciences* 14:471–517. [rKES]
Baron, J. (1998) *Judgment misguided: Intuition and error in public decision making.* Oxford University Press. [rKES]
Barwise, J. & Etchemendy, J. (1999) Language, proof, and logic. Seven Bridges Press. [SB]
(1994) *Hyperproof.* CSLI. [SB]
Bazerman, M., Baron, J. & Shonk, K. (2001) *"You can't enlarge the pie": Six barriers to effective government.* Basic Books. [rKES]
Belsky, G. & Gilovich, T. (1999) *Why smart people make big money mistakes – And how to correct them: Lessons from the new science of behavioral economics.* Simon & Schuster. [rKES]
Blackmore, S. (1999) *The meme machine.* Oxford University Press. [rKES]
Brakel, L. A. W. (in press a) Phantasies, neurotic-beliefs, and beliefs-proper. *American Journal of Psychoanalysis.* [LAWB]
(in press b) Phantasies and wishes: A proper function account of human a-rational primary process mediated mentation. *The Australasian Journal of Philosophy.* [LAWB]
Brakel, L. A. W., Kleinsorge, S., Snodgrass, J. M. & Shevrin, H. (2000) The primary process and the unconscious: Experimental evidence supporting two psychoanalytic presuppositions. *The International Journal of Psychoanalysis* 81:553–69. [LAWB]
Brakel, L. A. W., Villa, K. K. & Shevrin, H. (in preparation) Primary process categorization comes first: Experimental evidence supporting a psychoanalytic developmental hypothesis. [LAWB]
Bringsjord, S. (2000) In light of artificial intelligence, the science of mental ability is either silly or pointless. *Psycoloquy* 11(044). [SB]
Bringsjord, S., Bringsjord, E. & Noel, R. (1998) In defense of logical minds. *Proceedings of the 20th Annual Conference of the Cognitive Science Society,* pp. 173–78. Erlbaum. [SB]
Bringsjord, S. & Ferrucci, D. (1998) Logic and artificial intelligence: Divorced, separated, still married . . . ? *Minds and Machines* 8:273–308. [SB]
Carpenter, P., Just, M. & Shell, P. (1990) What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review* 97:404–31. [SB]
Cornford, F. M. (1945) Plato: *The republic,* trans. F. M. Cornford. Oxford University Press. [rKES]
Dawes, R. M. (2001) *Everyday irrationality.* Westview Press. [rKES]
Dennett, D. C. (1991) *Consciousness explained.* Little, Brown. [rKES]
Elio, R., ed. (2002) *Common sense, reasoning and rationality. Vancouver Studies in Cognitive Science, vol. 11.* Oxford University Press. [rKES]
Evans, J. St. B. T. (2002) Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin* 128:978–96. [rKES]
Evans, J. St. B. T. & Over, D. E. (1996) *Rationality and reasoning.* Psychology Press. [rKES]
Evans, J. St. B. T. & Wason, P. C. (1976) Rationalization in a reasoning task. *British Journal of Psychology* 67:479–86. [rKES]
Freud, A. (1936) *The ego and the mechanisms of defense.* International Universities Press. [LAWB]
Freud, S. (1895/1964) Project for a scientific psychology. In: *The standard edition of the complete psychological works of Sigmund Freud, vol. I,* ed. J. Strachey. Hogarth Press. [LAWB]
(1900/1964). *The Interpretation of Dreams.* In: *The standard edition of the complete psychological works of Sigmund Freud, vols. IV and V,* ed. J. Strachey. Hogarth Press. [LAWB]
(1911/1964) Formulation of the two principles of mental functioning. In: *The standard edition of the complete psychological works of Sigmund Freud, vol. XII,* ed. J. Strachey. Hogarth Press. [LAWB]
(1915/1964) The unconscious. In: *The standard edition of the complete psychological works of Sigmund Freud, vol. XIV,* ed. J. Strachey. Hogarth Press. [LAWB]
(1940/1964) An outline of psychoanalysis. In: *The standard edition of the complete psychological works of Sigmund Freud, vol. XXIII,* ed. J. Strachey. Hogarth Press. [LAWB]
Goel, V. & Dolan, R. J. (2003) Explaining modulation of reasoning by belief. *Cognition* 87:B11-B22. [rKES]
Johnson-Laird, P. (1983) *Mental models.* Harvard University Press. [SB]
Johnson-Laird, P. & Savary, F. (1995) How to make the impossible seem probable. *Proceedings of the 17th Annual Conference of the Cognitive Science Society,* pp. 381–84. Erlbaum. [SB]
Jou, J. (2000) Understanding/acceptance and adaptation: Is the non-normative thinking mode adaptive? *Behavioral and Brain Sciences* 23:680–81. [SB]
Kagel, C. J. (1987) Economics according to the rats (and pigeons too): What we have learned and what we hope to learn. In: *Laboratory experimentation in economics: Six points of view,* ed. A. Roth. Cambridge University Press. [rKES]
Kahneman, D. & Frederick, S. (2002) Representativeness revisited: Attribute substitution in intuitive judgment. In: *Heuristics and biases: The psychology of intuitive judgment,* ed. T. Gilovich, D. Griffin & D. Kahneman. Cambridge University Press. [rKES]
Kahneman, D. & Tversky, A., eds. (2000) *Choices, values, and frames.* Cambridge University Press. [rKES]
Klaczynski, P. A. (2001) Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development* 72:844–61. [rKES]
Klaczynski, P. A., Gordon, D. H. & Fauth, J. (1997) Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology* 89:470–85. [rKES]
Kokis, J., Macpherson, R., Toplak, M., West, R. F. & Stanovich, K. E. (2002) Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology* 83:26–52.
Kuhberger, A. (2002) The rationality of risky decisions: A changing message. *Theory and Psychology* 12:427–52.
Levelt, W. (1995) Chapters of psychology. In: *The science of the mind: 2001 and beyond,* ed. R. L. Solso & D. W. Massaro. Oxford University Press. [rKES]
Margolis, H. (1996) *Dealing with risk.* University of Chicago Press. [rKES]

Marr, D. (1982) *Vision.* W. H. Freeman. [rKES]
Millar, A. (2001) Rationality and higher-order intentionality. In: *Naturalism, evolution and mind*, ed. D. M. Walsh. Cambridge University Press. [rkES]
Moore, R. (1994) Making the transition to formal proof. *Educational Studies in Mathematics* 27:249–66. [SB]
Newell, A. (1982) The knowledge level. *Artificial Intelligence* 18:87–127. [rKES]
Newstead, S. E., Thompson, V. A. & Handley, S. J. (2002) Generating alternatives: A key component in human reasoning? *Memory and Cognition* 30:129–37. [rKES]
Newton, E. J. & Roberts, M. J. (2000) An experimental study on strategy development. *Memory and Cognition* 28:565–73. [EJN]
Oaksford, M. & Chater, N. (1995) Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning* 1:121–52. [rKES]
Oaksford, M. & Sellen, J. (2000) Paradoxical individual differences in conditional inference. *Behavioral and Brain Sciences* 23:691–92. [rkES]
Over, D. E. (2000) Ecological rationality and its heuristics. *Thinking and Reasoning* 6:182–92. [rKES]
(2002) The rationality of evolutionary psychology. In: *Reason and nature: Essays in the theory of rationality,* ed. J. L. Bermudez & A. Millar. Oxford University Press. [rKES]
Perkins, D. N. (1995) *Outsmarting IQ: The emerging science of learnable intelligence.* Free Press. [rKES]
Real, L. A. (1991) Animal choice behavior and the evolution of cognitive architecture. *Science* 253:980–86. [rKES]
Rips. L. (1994) *The psychology of proof.* MIT Press. [SB]
Roberts, M. J., Gilmore, D. J. & Wood, D. J. (1997) Individual differences and strategy selection in reasoning. *British Journal of Psychology* 88:473–92. [EJN]
Russo, J. E. & Schoemaker, P. (2002) *Winning decisions: Getting it right the first time.* Doubleday. [rKES]
Sa, W., West, R. F. & Stanovich, K. E. (1999) The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology* 91:497–510. [rKES]
Schneider, S. (2000) An elitist naturalistic fallacy and the automatic controlled continuum. *Behavioral and Brain Sciences* 23:695–96. [SB]
Schommer, M. (1990) Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology* 82:498–504. [rKES]
Shafir, E., Simonson, I. & Tversky, A. (1993) Reason-based choice. *Cognition* 49:11–36. [rKES]
Shafir, E., Waite, T. A. & Smith, B. H. (2002) Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology* 51:180–87. [rKES]
Shiffrin, R. M. & Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127–90. [rKES]
Sinatra, G. M. & Pintrich, P. R., eds. (2003) *Intentional conceptual change.* Erlbaum. [rKES]
Sloman, S. A. (2002) Two systems of reasoning. In: *Heuristics and biases: The psychology of intuitive judgment,* ed. T. Gilovich, D. Griffin & D. Kahneman. Cambridge University Press. [rKES]
Sloman, S. A. & Rips, L. J., eds. (1998) *Similarity and symbols in human thinking.* MIT Press. [rKES]
Slovic, P., Finucane, M. L., Peters, E. & MacGregor, D. G. (2002) The affect heuristic. In: *Heuristics and biases: The psychology of intuitive judgment,* ed. T. Gilovich, D. Griffin & D. Kahneman. Cambridge University Press. [rKES]
Southerland, S. A. & Sinatra, G. M. (2003) Learning about biological evolution: A special case of intentional conceptual change. In: *Intentional conceptual change,* ed. G. M. Sinatra & P. R. Pintrich. Erlbaum. [rKES]
Stanovich, K. E. (1999) *Who is rational? Studies of individual differences in reasoning.* Erlbaum. [rKES]
(2000) *Progress in understanding reading: Scientific foundations and new frontiers.* Guilford Press. [rKES]
(2004) *The robot's rebellion: Finding meaning in the age of Darwin.* University of Chicago Press. [rKES]
Stanovich, K. E. & West, R. F. (1997) Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology* 89:342–57. [rKES]
(2003) Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. In: *Evolution and the psychology of thinking: The debate,* ed. D. Over. Psychology Press. [rKES]
Sternberg, R. J. (1997) *Thinking styles.* Cambridge University Press. [rKES]
(2000) The ability is not general, and neither are the conclusions. *Behavioral and Brain Sciences* 23:697–98. [SB]
Sternberg, R. J., ed. (2002) *Why smart people can be so stupid.* Yale University Press. [rKES]
Todd, P. M. & Gigerenzer, G. (2000) Précis of *Simple heuristics that make us smart. Behavioral and Brain Sciences* 23:727–80. [rKES]
Toplak, M. & Stanovich, K. E. (2002) The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology* 94:197–209. [rKES]
West, R. F. & Stanovich, K. E. (1978) Automatic contextual facilitation in readers of three ages. *Child Development* 49:717–27. [rKES]
(1986) Robust effects of syntactic structure on visual word processing. *Memory and Cognition* 14:104–12. [rKES]
Yang, Y. & Bringsjord, S. (under review) Mental metalogic: A new paradigm in psychology of reasoning. [SB]
Zizzo, D. (2000) Implicit learning of (boundedly) rational behavior. *Behavioral and Brain Sciences* 23: 700–701. [SB]