AJP

REFLECTIONS ON RATIONALITY

Rationality and the Reflective Mind

By Keith E. Stanovich. New York, NY: Oxford University Press, 2011. 328 pp. Hardcover, $65.

Keith Stanovich is an outstanding scholar and cognitive scientist who has received many awards in recognition of his contributions to psychology and education. His decision, together with his regular collaborator, Rich West, to move their program of work into the psychology of reasoning and decision making from the 1990s onward has been of great benefit to these research fields and has led to numerous journal articles and several books before this one. In particular, their large-scale psychometric studies of individual differences in cognitive tasks have added a new dimension of understanding of the cognitive processes involved. At the theoretical level, Stanovich has been at the forefront of developing the dual process theory of reasoning and judgment and its underlying cognitive architecture. This new book represents the substantial changes and theoretical developments he has made over the past few years.

Although he has written other books in between, *Rationality and the Reflective Mind* is best viewed as a sequel to Stanovich's well-known earlier book, *Who Is Rational?* (1999). It deals with essentially the same issues: rationality, dual process theory, and individual differences in performance on reasoning and decision-making tasks. Research on all these topics has moved on considerably in the meantime, but so has Stanovich. The advances in his thinking about cognitive architecture are particularly impressive, as is his unrivaled scholarship. Quite apart from the new theoretical thinking, *Rationality and the Reflective Mind* provides a comprehensive and unrivaled source of academic references on a very wide range of topics in higher cognition and the study of intelligence. However, Stanovich has wisely consigned many of these to footnotes so as to keep the main text as readable as possible.

Although there has been no formal collaboration between Stanovich and me, our work has been mutually influential over the past 15 years or so. Readers of my own book *Thinking Twice* (2010), which Stanovich reviews in tandem with this one, will note a number of similarities in our thinking. However, there are significant differences in our research programs also. Although I focus here mostly on what Stanovich has to say, I will note also these similarities and differences. In fact, Stanovich's research program

uniquely combines three main strands, all of which are well developed and presented in his new book. These strands concern his approach to rationality, dual processing and cognitive architecture, and individual differences in performance. The similarities to my own work lie largely in dual processing and cognitive architecture. I have conducted little research on individual differences, and it has had less influence on my thinking about architecture. And our views on rationality are somewhat different. I begin with this topic, as indeed does Stanovich in his opening chapter.

Stanovich focuses mostly on instrumental rationality (attaining one's goals), which he links with certain forms of normative theory, such as Bayesian decision theory. Hence, he regards reasoning and judgment tasks as having right and wrong answers. He is careful to point out that rationality can occur only at the personal level. Subpersonal systems, such as the visual system or episodic memory, can only vary in efficiency; they cannot be *irrational*. The link with dual process theory comes in the distinction between automated processes and reflective and volitional forms of cognition. Although he famously dubbed these as Systems 1 and 2 in his 1999 book, he now (like me) prefers to avoid these terms, which may tempt readers to think that there are just two brain systems responsible for these two kinds of processing. Instead, he contrasts type 1, or heuristic, processes with type 2, or analytic, processes. Only at the type 2 level can we talk about rationality rather than efficiency. He links all this with a tripartite theory of the mind, which I discuss later.

Stanovich portrays a Great Rationality Debate as primarily between two camps: Meliorists and Panglossians. Stanovich himself is a self-declared Meliorist, meaning that he thinks people can have degrees of rationality that can be improved by experience and training. For this he needs normative standards: It must be possible to say that some people are performing better than others on cognitive tasks. Panglossians, in which he includes philosophers such as Cohen (1981) and the bulk of evolutionary psychologists, appear to propose that all behavior is adaptive and therefore rational by definition. Hence, normative theories are constructed post hoc according to the behavior observed. However, there is another position that falls outside this debate: One can reject normative analysis in favor of an essentially descriptive approach to the psychology of thinking (Elqayam & Evans, 2011). For example, my own recent review of higher cognition, framed around hypothetical thinking theory (Evans, 2007), attempted

to minimize reference to any normative theory of behavior. However, Stanovich and I agree on the essentially instrumental nature of behavior: The mind must be understood in terms of the goals people pursue and the mechanisms that have evolved for this purpose. And the specific dual process accounts we each give of particular cognitive tasks are broadly similar.

The second strand in Stanovich's work is his development of dual process theory into a tripartite model of the mind. What was "System 1" or "the set of autonomous subsystems" (Stanovich, 2004) is now the *autonomous mind*. What was previously System 2 has been bifurcated into the *algorithmic* and *reflective minds*. The algorithmic mind provides the essential computational machinery for type 2 thinking, particularly the ability to decouple suppositions from beliefs and engage in hypothetical thinking. The reflective mind includes the goals, disposition, and strategies that people adopt in their type 2 processing. This seems to be the personal level capable of being rational or irrational in his theory. My own latest scheme is a two (not three) mind theory laid out in *Thinking Twice*. However, the difference between us is less dramatic than it may appear. What I call the old or intuitive mind corresponds pretty much with Stanovich's autonomous mind. My new or reflective mind incorporates both his algorithmic and reflective levels. I put a lot of emphasis on engagement of central working memory (or controlled attention) as a defining feature of type 2 processing; this facility seems to correspond fairly closely to Stanovich's algorithmic mind. He measures the efficiency of this by individual differences in general intelligence, but these are known to be very highly correlated with measures of working memory capacity (Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004). He also puts a lot of emphasis on what he calls "decoupling" in type 2 processes: the ability to separate supposition from belief in hypothetical thinking. These facilities also have featured strongly in my own writing about the nature and purpose of type 2 thinking (Evans, 2007; Evans & Over, 1996).

To understand the rest of his cognitive architecture, we need to appreciate Stanovich's third strand: the study of individual differences in intelligence and rational thinking. For several years now, he has been emphasizing a very important point: There is more to being rational than having a high IQ (Stanovich, 2009). IQ tests are a measure cognitive capacity and hence the efficiency of what he calls the algorithmic mind. However, rational thinking requires a number of other things, which is the main focus of this book. Instrumental rationality (achievement of goals) is de-

pendent on *epistemic* rationality (the acquisition of true and relevant beliefs). For example, if people hold false beliefs or pursue inappropriate goals, then we will judge their behavior to be irrational, no matter how high their IQ. Stanovich has coined the term *mindware* to refer to the beliefs and procedures people need to acquire (e.g., by education) for rational thought in different domains. Mastering the integral and differential calculi, for example, provides essential mindware for a mathematician. Also of great importance to him are individual differences in *rational thinking dispositions*. Regardless of their ability, some people are more inclined to apply analytic thinking to problems, whereas others rely on intuition. According to Stanovich, intuitive processing often leads to cognitive biases, especially in modern technological environments, which are very different from the environment of evolutionary adaptation. One of his running arguments with evolutionary psychology is to contest their assumption that type 1 processing is generally adaptive in the modern world.

Although I understand Stanovich's reasons for dividing the new mind into algorithmic and reflective components, I have some reservations about this. I think it is driven by the psychometric perspective that features so strongly in his work. It is certainly true that there are reliable individual differences in both cognitive ability (IQ or working memory capacity) on one hand and rational thinking dispositions on the other. As he demonstrates in this book, the two have a low correlation with each other, and each accounts for separate parts of the variance associated with individual differences in performance on reasoning and judgment tasks. It is hard to argue with his critique of intelligence testing and the critical point that IQ is at most necessary and never sufficient for rational thinking. One of the key developments since his 1999 book is the demonstration that a number of cognitive biases (e.g., myside bias) operate almost independently of IQ. None of this to my mind makes the case that the algorithmic and reflective minds are architecturally distinct entities as cognitive or brain systems. I just see them as two distinct properties of the new mind. However, this difference is more a matter of semantics than substance. As Stanovich points out, I have proposed the operation of type 3 higher-level control processes (Evans, 2009) that perform a similar function to his reflective mind.

Having outlined the three main strands that drive Stanovich's work and the broad manner in which his theory has developed, I need to say more specifically about what this book provides. The short answer is

a very great deal. First, even for those familiar with his frequent publications, Stanovich provides discussion here of a number of concepts that are generally better developed and integrated than in previous writing, and several will be new to many of his readers. One example of work that was new to me was an interesting critical discussion of measures of executive function, on which there is large literature spun off from the working memory movement. It is important to Stanovich to show that these tap into the algorithmic mind (capacity) and not the reflective mind (disposition), and so he argues that the term *executive function* is misleading and this work often misinterpreted. Another example is the discussion of studies in neuroscience and especially the role of the anterior cingulate cortex in detecting conflict and regulating cognitive control.

Stanovich's thinking about cognitive biases has moved on or been clarified in important ways. His 1999 book and some of my earlier writing (Evans, 1989; Evans & Over, 1996) may have contributed to a fallacious belief that type 1 processing = bias and type 2 processing = normatively correct. This was never actually our position, but the emphasis given in earlier writing encouraged this simplistic view, which I have been arguing against in recent publications (e.g., Evans, 2007, 2008). For example, the psychometric studies discussed by Stanovich (1999) predominantly showed that normative solutions were associated with those of higher general intelligence. But one of the major developments in his research program (with Rich West) since that time has been the demonstration of many cognitive biases for which IQ offers little protection (e.g., Stanovich & West, 2008). And of course, his central argument is that measured intelligence is never sufficient for rational thinking. Hence, he very clearly states in this book that both type 1 and 2 processing can lead to correct answers, but each also can lead to biases. Another key development here is that he proposes a particular form of low-effort and generally ineffective type 2 thinking, which he calls *serial associative cognition*.

Much emphasis in the book is given to the idea that for good evolutionary reasons, people are cognitive misers, who by default rely on either type 1 intuitive processing or low-effort type 2 thinking (serial associative cognition). Their precious and limited resources for decoupling and mental simulation (what I would call working memory) are applied sparingly. Thus, the detailed model of reasoning errors that he presents includes several factors independent of intelligence. First, a person must detect the need for intervention with type 2 thinking; they must also be inclined to intervene, which is where rational thinking dispositions come in as a personality characteristic; if (and only if) success requires decoupling and mental simulation, then the capacity of the algorithmic mind is relevant, and IQ (or working memory capacity) becomes a predictor of performance. In addition, errors of type 2 thinking may result from missing or contaminated mindware. This is a complex model with more parameters than one would ideally like, but it does seem justified by a detailed study of the literature. And no one studies it in more detail than Stanovich. An impressive feature of the later chapters is some very lengthy and detailed tables classifying a range of reasoning, decision, and other rational thinking tasks (e.g., from the study of social cognition). In essence, Stanovich has analyzed and explicitly stated the nature of the biases and errors in his framework for practically every relevant cognitive task in the psychological literature. This is supported in his detailed taxonomies of biases and in his report of some new empirical work on heuristics and bias tasks. The scholarship and level of analytic detail shown are hugely impressive; no other scholar has come close to this level of mastery of these literatures.

Although he has written before about higher-order preference and values that feature uniquely in human cognition (e.g., Stanovich, 2004), these ideas now seem better developed and integrated with his dual process framework. He has recently coined the term *Master Rationality Motive (MRM)* and devotes a whole chapter of the book to its discussion. This motive is possible only because of the uniquely human capacity for meta-representation, the ability to represent and reflect on the beliefs of ourselves and others (see also *Thinking Twice*). Stanovich links MRM with the evolution of type 2 processing and unique capacity of humans to be rational in the personal sense. They strive for rational integration. Specifically, the MRM is the desire to behave according to one's own higher-order values and principles. For example, there are three kinds of smokers. First, there are those who just want to smoke and do not reflect on it (Stanovich calls these *wantons*). Then there are those who want not to want to smoke, but they smoke anyway because of the addiction. Finally, there are those who want to want to smoke. For example, they may argue that the benefits of smoking (pleasure, relaxation, weight loss) outweigh its long-term health risks. Perhaps only the second of these groups can be said to be irrational in a personal sense. In their two minds conflict, the struggle for control is won at the impersonal level of the old mind.

The book concludes with two chapters that discuss intelligence and rationality, capturing the essence of his purpose in writing this volume. He is clearly very unhappy with the increasingly broad definitions of intelligence, which get easily confused with the narrow definition used to measure IQ. This leads to false aggrandizement of IQ, which is a highly incomplete measure of rational thought. As a result, "we have been valuing only the algorithmic mind and not the reflective mind" (p. 187). The final chapter presents us with yet another detailed taxonomy of rational thinking skills and develops that argument that we need to find an "RQ," a test of rational thinking that avoids the inadequacies of the IQ test. The final chapters do not hesitate to spell out the social and political implications of these arguments.

In conclusion, *Rationality and the Reflective Mind* is an extraordinary achievement, reflecting 15 years or so of extensive research, detailed scholarship, and the clearest theoretical thinking about the psychology of higher cognitive processes. The list of people who should read and study it is long: At the very least it should include all psychologists engaged in empirical study of thinking, reasoning, and decision making; all cognitive scientists, neuroscientists, and philosophers interested in the cognitive architecture of the human mind; and all psychologists and educationalists interested in the measurement of intelligence and thinking dispositions and their application in society. At their best, academic books can provide a level of development and integration that is not possible with the constraints of individual journal articles. This one certainly does.

Jonathan St. B. T. Evans
University of Plymouth
School of Psychology
Plymouth PL4 8AA
England, UK
E-mail: j.evans@plymouth.ac.uk

REFERENCES

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4,* 317–370.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence, 32,* 277–296.

Elqayam, S., & Evans, J. St. B. T. (2011). Subtracting "ought" from "is": Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences, 34,* 233–290.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hove, UK: Erlbaum.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement.* Hove, UK: Psychology Press.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology, 59,* 255–278.

Evans, J. St. B. T. (2009). How many dual-process theories do we need: One, two or many? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 31–54). Oxford, England: Oxford University Press.

Evans, J. St. B. T. (2010). *Thinking twice: Two minds in one brain.* Oxford, England: Oxford University Press.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, UK: Psychology Press.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin.* Chicago, IL: Chicago University Press.

Stanovich, K. E. (2009). *What intelligence tests miss. The psychology of rational thought.* New Haven, CT: Yale University Press.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672–695.

CUMULATIVE PROGRESS IN UNDERSTANDING OUR MULTIPLE MINDS

Thinking Twice: Two Minds in One Brain
By Jonathan St. B. T. Evans. Oxford, England: Oxford University Press, 2010. 240 pp. Cloth, $59.95.

The notion that our brains carry out qualitatively different types of processing—in essence, that we have many different minds within our brains—is not new. Tracing back to Plato, the idea has recurred in a variety of disciplines and has been around for many decades in the modern period of psychology. What is new, however, is that cognitive scientists have begun to understand the biology and cognitive structure of these qualitatively different types of processing. Much of our current understanding is embodied in the so-called dual process theories of cognition that have received a large share of attention in the last decade (e.g., Evans & Frankish, 2009; Kahneman & Frederick, 2002; Lieberman, 2007). No one is better placed to take us on a tour of these recent developments than is Jonathan Evans. With Peter Wason, Ev-

ans was one of the first psychologists in the modern era to posit a scientifically plausible dual process idea (Wason & Evans, 1975). When the idea went fallow in the 1980s, Evans (1984, 1989) continued to explore its usefulness. And when dual process theory's current popularity was ignited in the late 1990s (Evans, 2003; Kahneman & Frederick, 2002; Sloman, 1996; Stanovich, 1999), it was Jonathan Evans who had been tending the flame for two decades.

In *Thinking Twice: Two Minds in One Brain,* Evans has written one of those extraordinary books that can speak to two audiences at once. This book is a fascinating layperson's introduction to the implications of dual process theory across an astonishing range of psychological specialties, from the reasoning and decision-making literature, to neuropsychology, clinical psychology, and consciousness. But the book is not just for the layperson. Indeed, for the active researcher it contains one of the best overall contextualizations of the state of dual process theory. I shall emphasize the contributions for the latter audience, given where this review appears.

Dual process theory has been the subject of many critiques that treat the theory as a straw man by presenting it in extreme or dated form (see Stanovich & Toplak, in press). Because, as outlined earlier, the bare bones of the theory have been around since the 1970s, it is a fat target to attack a dated version containing features that have long since been superseded. Evans's book is particularly important in presenting a contemporary version, free of the encumbrances of long-abandoned assumptions. For example, Evans uses the terms Type 1 and Type 2 processing to indicate that he is discussing a dual process theory, not a dual system one. In this and many other ways, Evans provides us with a contemporary dual process view, not one frozen static by some seminal articles now almost 15 years old.

Evans is particularly good at dissecting the recent enthusiasm for intuition and nonconscious processing both in psychology and in the general media. He picks out the obvious example of Malcolm Gladwell's book *Blink* (2005) from the general media and focuses on Gigerenzer's work as an example of the former. Evans is spot on when he argues that "these authors claim, or come very close to claiming, that intuition is king and that we are better off not trying to second guess its powers with conscious reasoning" (p. 94). He humorously refers to their stance as the "no-mind position." Evans then goes on to specifically show the flaws in the claim that Type 2 (analytic) processing is almost superfluous given the enormous powers of Type 1 processing and the autonomous mind.

Many of Gladwell's examples rest on an old finding in psychology: that many complex rules, stimulus discriminations, and decision-making principles can be practiced to automaticity so that they eventually become processed in a Type 1 manner. Thus, tightly compiled learned information that becomes executable in a Type 1 manner has been a crucial component of dual process models from the beginning (e.g., Shiffrin & Schneider, 1977). This is relevant because the no-mind theorists exploit the fact that Type 1 processing is multifarious and does not arise from a singular system (a point made by Evans in several of his previous writings). The many kinds of Type 1 processing have in common the property of autonomy, but otherwise their neurophysiology and etiology might be very different. For example, Type 1 processing would include behavioral regulation by the emotions, the encapsulated modules for solving specific adaptive problems that have been posited by evolutionary psychologists, processes of implicit learning, and the automatic firing of overlearned associations.

Evans explains how the no-mind theorists exploit the ambiguity in the term *Type 1 processing* or *intuitive processing*. They imply that such processing and the feats it accomplishes are not at all attributable to Type 2 processing, or what is sometimes called the reflective mind. But of course overlearned associations or discriminations that come to be made in a Type 1 manner are nothing of the sort. The reflective mind was involved from the very beginning—when the training was conscious, when it was planned, and when the motivational force of the reflective mind was marshaled in order to sustain the practice. The art experts in the no-mind literature who make complex intuitive judgments are making those judgments because of a much longer period of reflective thinking that has gone on before.

Evans's discussion of Gigerenzer's work is even more nuanced because the conceptual mistakes being made by the fast and frugal tradition are more complex. A simple example will give a flavor of what is going wrong with that version of the no-mind position. The recognition heuristic studied by the Gigerenzer group is well known to many researchers. It is often classified as an example of Type 1 processing, or as the product of the intuitive mind. But that cannot be right, if one starts to think about the complexity of its use in actual situations rather than those just in the laboratory. Evans posits, "Suppose I am asked to judge the relative size of German cities. Doubtless, I will use

the recognition heuristic for most of these, and judge more familiar sounding cities to be larger. Suppose, however, the experimenter mentions a small city, which has high recognition value for me. Perhaps I visited it on a recent holiday. I won't say it is a large city (because I recognize it) but rather use my actual knowledge that it is *small* to make the opposite judgment. This shows that the recognition heuristic is under full control of the reflective mind and not being applied *mindlessly*" (p. 99). Evans's summary of the work of the advocates of fast and frugal heuristics is critical but eminently fair. He acknowledges that quickly applied heuristics may well be superior to extended analytic processing in some cases, but only when we have induced from experience much relevant information and interfaced it with the proper responses.

Evans's book does not stop with the traditional topics in thinking and reasoning, however. He takes us on a whirlwind tour of how dual process theory meshes with the problems of social psychology and with the issues surrounding consciousness and free will. Evans's own hypothetical thinking theory has much to say about the role of imagination in human mental life. Applied problems such as the treatment of phobias and pathological gambling make their appearance and are illuminated by understanding how they fit within the two-minds view.

Of particular interest to me, given my own interest in the psychology of rational thought (Stanovich, 2004, 2009), is that Evans gets issues related to rationality right. He baldly states that "there can be no rationality *without* emotion" (p. 187), thus contradicting the popular but incorrect view in folk psychology that emotion is antithetical to rationality. This common idea sees the absence of emotion as purifying thinking into purely rational form. This idea is not consistent with the definition of rationality in modern cognitive science. Instrumental rationality is behavior consistent with maximizing goal satisfaction, not a particular psychological process. It is perfectly possible for the emotions to facilitate instrumental rationality as well as to impede it. In fact, conceptions of emotions in cognitive science stress the adaptive regulatory powers of the emotions. The basic idea is that emotions stop the combinatorial explosion of possibilities that would occur if an intelligent system tried to calculate the utility of all possible future outcomes. Emotions are thought to constrain the possibilities to a manageable number based on similar situations in the past.

In short, emotions get us in the right ballpark of the correct response. If more accuracy than that is required, then a more precise type of analytic cognition will be needed. Of course, we can rely too much on the emotions. We can base responses on a ballpark solution in situations that really need a more precise type of analytic thought. More often than not, however, processes of emotional regulation facilitate rational thought and action.

In parts of the same chapter on dual process theory and rationality, Evans gets a crucial distinction just right: the distinction between (subpersonal) genetic goals and the goals of the individual person. This distinction lies behind the presumption (see Stanovich, 2004) that the statistical distributions of the types of goals being pursued by Type 1 and Type 2 processing might be different and that important consequences for human self-fulfillment follow from this fact. The greater evolutionary age of some of the mechanisms underlying Type 1 processing accounts for why it more closely tracks ancient evolutionary goals (i.e., the genes' goals), whereas Type 2 processing instantiates a more flexible goal hierarchy that is oriented toward maximizing overall goal satisfaction at the level of the whole organism. Because Type 2 processing is more attuned to the person's needs as a coherent organism than is Type 1 processing, in the minority of cases where the outputs of the two systems conflict, people will often be better off if they can accomplish a system override of the Type 1–triggered output.

Evans's essential endorsement of this conjecture reinforces a criticism of some work in evolutionary psychology and adaptive modeling for implicitly undervaluing instrumental rationality by defending nonnormative responses made by many subjects in reasoning experiments. Many such instances occur when there is a conflict between the responses primed by Type 1 and Type 2 processing and the former dominates. Such situations are interpreted within a dual process framework as reflecting conflicts between two different types of optimization: fitness maximization at the subpersonal genetic level and utility maximization at the personal level. Evolutionarily adaptive behavior is not the same as rational behavior. Evolutionary psychologists obscure this by sometimes implying that if a behavior is adaptive, it is rational. Such a conflation represents a fundamental error of much import for human affairs. Definitions of rationality must be kept consistent with the entity whose optimization is at issue. In order to maintain this consistency, the different "interests" of the replicators and the vehicle must be explicitly recognized (Dawkins, 1976/1989; Dennett, 1995).

A failure to differentiate these interests is at the heart of the disputes between researchers working in the heuristics and biases tradition and their critics in the evolutionary psychology camp (for discussions of these debates, see Kahneman & Tversky, 1996; Samuels & Stich, 2004; Stanovich, 1999, 2004; Stein, 1996). The empirical demonstrations of a gap between descriptive and normative models of reasoning have been reinterpreted by various evolutionary psychologists, adaptationist modelers, and ecological theorists. These theorists have reinterpreted the modal response in most of the classic heuristics and biases experiments as indicating an optimal information processing adaptation on the part of the subjects. In the extreme, these investigators have argued that the research in the heuristics and biases tradition has not demonstrated any human irrationality at all. The evolutionary psychologists are probably correct that most Type 1 processing is evolutionarily adaptive. What they have failed to realize is that their evolutionary interpretations do not impeach the position of the heuristics and biases researchers that the alternative response given by the minority of subjects is rational at the level of the individual.

Evans includes a chapter on the two minds in conflict (and in cooperation). The idea of minds in conflict finds motivation from an observation that is consistent with all the classes of dual process theories that Evans discusses: Many people making poor choices are alienated from the choices they make. For example, Reyna and Farley (2006) point out that people "who take unhealthy risks often agree that their behavior is irrational, on sober reflection, but they gave in to temptation or were not thinking at the time of the decision, and are worse off for having done so" (p. 35). Instead of the economics-like assumption of people as coherent rational actors, "multiple minds" theories, from Minsky (1985) and Dennett (1996) to their contemporary incarnations (e.g., Evans, 2003, and the present book), highlight the image of a decision maker in conflict. This comports well with the fact that many people with behavioral problems (the examples of phobias, gambling, and other addictions are all discussed in Evans's book) will indeed verbally reject their own behavior. This second-order judgment (to desire to desire differently) is a critical component of some dual process views and is something that can be built on therapeutically to motivate cognitive reform (see Stanovich, 2004).

Throughout the book, Evans does not shrink from large conclusions, nor does he flinch when de-livering conclusions that might startle the lay reader. Consider the following passage near the end of the book: "You may have a different view of yourself than you had before picking up this book. Like most people, you may have assumed the common sense, folk psychological view of the mind in which a conscious person, a chief executive, sits at the top in control. The psychological and neuroscientific evidence discussed in this book overwhelmingly refutes this" (p. 207). This is one of many places in the book where Evans echoes the themes of Dennett (1991, 1996, 2003) and conveys them to the layperson with remarkable economy of expression. The lay reader may not warm to all these Dennett-like themes, but Evans explains clearly how they follow from contemporary work in cognitive science.

Another theme in this book that echoes Dennett (1996) is the stress placed on the fact that the evolutionarily more recent minds did not *replace* earlier ones but instead were added to them. In developmental psychology, the mistaken idea that Type 2 processing replaces Type 1 processing in the brain is called the "illusion of replacement" (Brainerd & Reyna, 2001, p. 52). Evans is at pains to show that the evolutionary story does not play out in this manner. In doing so, he explains to the lay reader why humans seem to be similar to other animals in some ways but not others. The similarities stem from the old mind we share with other animals and the differences from the new mind that Evans describes as not unique to humans but uniquely *developed* in humans.

Especially interesting for researchers is the Addendum to the book, which deals with some technical issues that might not be of as much interest to the lay reader. However, all are of great interest to researchers concerned with the nuances of dual process theories and the latest developments in their evolution. To take one example from the Addendum, Evans discusses the difference between types and modes of processing. Confusing the two has been common when discussing correlations of task performance with thinking dispositions. Thinking dispositions are not associated with Type 1 or Type 2 processing, as implied in some writings. For example, implying that it signals some kind of inconsistency in dual process views, Newstead (2000) argues that "Epstein, Pacini, Denes-Raj & Heier (1996) found that superstitious and categorical thinking, which might be supposed to be part of System 1, produced no significant correlations, either positive or negative, with Faith in Intuition (System 1)" (p. 690). But superstitious thinking signals a mode of thought, not a type, and this disposition is not at

all an indicator of the functioning of Type 1 process-ing. Because Type 2 processing is the only type of processing that is characterized by flexible goals and flexible cognitive control, it is variation in this type of processing that *all* thinking disposition measures are assessing. Modes (thinking dispositions) are not associated with one system or the other; they assess differences in the style of Type 2 processing. As I have previously discussed (Stanovich, 1999, 2009), psy-chometric intelligence relates to the algorithmic level of Type 2 processing, whereas thinking dispositions such as superstitious thinking, need for cognition, and actively open-minded thinking relate to styles of epis-temic and goal regulation at the reflective level of Type 2 processing. Newstead's mistake is a common one, so the distinction between modes and types of thinking that Evans discusses is important. In the Addendum, Evans treats a variety of other important issues with a scholarly care that will be much appreciated by the academic audience of the book.

Versions of dual process theory are now echoing throughout cognitive psychology, social psychology, clinical psychology, and neuropsychology. This book is ideally placed to capture this interest. The author has contributed both empirically and theoretically to the development of these models, and his book provides very nice and balanced coverage of both the reasoning literature and the decision-making literature. Dual pro-cess theorists often segregate themselves into one or the other of these academic literatures, so the coverage here is most welcome. The reader gains from an author immersed in both of these literatures. Broad in scope and appropriate for a wide range of audiences—from lay readers to undergraduates to researchers—this book is the perfect guide for newcomers to the dual process literature and for those wanting an accurate update on recent trends.

Keith E. Stanovich
Department of Human Development and Applied Psy-chology
University of Toronto
252 Bloor St. West
Toronto, Ontario
Canada M5S 1V6
E-mail: keith.stanovich@utoronto.ca

REFERENCES

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neu-roscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, pp. 41–100). San Diego, CA: Academic Press.

Dawkins, R. (1989). *The selfish gene.* New York, NY: Oxford University Press. (Original work published 1976)

Dennett, D. C. (1991). *Consciousness explained.* Boston, MA: Little Brown.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life.* New York, NY: Simon & Schuster.

Dennett, D. C. (1996). *Kinds of minds: Toward an under-standing of consciousness.* New York, NY: Basic Books.

Dennett, D. C. (2003). *Freedom evolves.* New York, NY: Vi-king.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of Personality and Social Psychology, 71,* 390–405.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75,* 451–468.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hove, UK: Erlbaum.

Evans, J. St. B. T. (2003). In two minds: Dual-process ac-counts of reasoning. *Trends in Cognitive Sciences, 7,* 454–459.

Evans, J. St. B. T., & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond.* Oxford, England: Ox-ford University Press.

Gladwell, M. (2005). *Blink.* New York: Little, Brown.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuris-tics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Kahneman, D., & Tversky, A. (1996). On the reality of cogni-tive illusions. *Psychological Review, 103,* 582–591.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology, 58,* 259–289.

Minsky, M. (1985). *The society of mind.* New York, NY: Simon & Schuster.

Newstead, S. E. (2000). Are there two different types of thinking? *Behavioral and Brain Sciences, 23,* 690–691.

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making. *Psychological Science in the Public Interest, 7,* 1–44.

Samuels, R., & Stich, S. P. (2004). Rationality and psychol-ogy. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 279–300). Oxford, England: Oxford University Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psy-chological Review, 84,* 127–190.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Stanovich, K. E. (1999). *Who is rational? Studies of individu-al differences in reasoning.* Mahwah, NJ: Erlbaum.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago, IL: University of Chicago Press.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K. E., & Toplak, M. E. (in press). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind & Society*.

Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford, England: Oxford University Press.

Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition, 3*, 141–154.

SELIGMAN'S FLOURISH: THE SECOND COMING

Flourish: A Visionary New Understanding of Happiness and Well-Being

By Martin E. P. Seligman. New York, NY: Free Press, 2011. 368 pp. Hardcover, $26.

During his presidency of the American Psychological Association in 1998, Martin Seligman proposed that psychological science focus less on pathology and repair and more on the realization of human potential. Today, "positive psychology" is a well-known movement within the field. Seligman's career began with the study of rather negative experiences. His graduate work emerged from the fear conditioning studies conducted in Richard Solomon's laboratory (Solomon & Wynne, 1953). Dogs were subjected to inescapable electric shock until they failed to learn how to leap to safety when opportunities to do so became available (Overmier & Leaf, 1965; Overmier & Seligman, 1967).

The theory of "learned helplessness"[1] and its reframing in terms of Weiner's (1974) attribution theory (Abramson, Seligman, & Teasdale, 1978) made Seligman's name in clinical and social psychology. In time, Seligman set his sights on the other side of the coin. If there is learned helplessness, there must also be learned optimism (Seligman, 1998). This perspective shift was a metaphorical play on Solomon's opponent-process model of emotion (Solomon & Corbit, 1974). Each emotion, once triggered, eventually brings along its opposite. Where there is despair, there shall also be hope.

I remember the beginning of positive psychology in January 1999. Martin Seligman, Mihaly Csikszentmihalyi, and Ray Fowler had organized an invitation-only conference in Akumal in the Yucatán. The attendees were about two dozen youngish psychologists curious about the positive psychology initiative. Some of them are still associated with the movement today.

Positive psychology is not egalitarian in theory or in practice. In Akumal, the differences in seniority and status were evident. The attendees were advised not to bring their families, and they had to sleep in bunk beds; the conveners had posher digs, brought along family, and commuted in Range Rovers. Hierarchy seems to come naturally in our society, and it highlights a conceptual difficulty in positive psychology. On one hand, positive psychology emphasizes individual differences in character strengths, virtues, and accomplishments. It endorses a psychological meritocracy. For anyone who is strong on a positive character trait, someone else must be low. It is not possible for everyone to be distinguished. On the other hand, positive psychology emphasizes the overall, societal level of desirable character traits. Seligman argues that his grand objective is to improve character and virtue across the board. However, average levels in a group and differences between individuals are independent only in theory. If there is an upper limit—as there usually is with anything that humans can be or can accomplish—individual differences constrain changes in level. Conversely, if everyone reached the maximum of human potential, no one would be distinguished or even be distinguishable.

Why the new book? Seligman writes that he has changed his mind about positive psychology. He presents *Flourish: A Visionary New Understanding of Happiness and Well-Being* as a radical departure from his earlier view described in *Authentic Happiness* (Seligman, 2002). He has abandoned the belief that positivity is all about happiness. Happiness is a good target for self-report measures because it is so subjective (Gilbert, 2006). But measurement errors may occur. Some people who think they are happy really are not, and only trained observers can tell the difference (Shedler, Mayman, & Manis, 1993). In *Flourish,* Seligman also considers the obverse, people who are actually more well than they think they are. He suggests counting one's blessings as an exercise to calibrate subjective judgment of our happiness levels.

In *Authentic Happiness,* Seligman proposed positive affect, engagement with tractable tasks, and meaning as the building blocks of happiness. Using a spiritual definition of *meaning,* Seligman appeals to a sense of connection with something greater than the self. This could mean a lot of different things to

different people.[2] In *Flourish,* Seligman adds positive social relationships and accomplishments as criteria. Students of Occam learn that a theory should be allowed to take on new elements only when necessary, lest the theory lose conciseness and power. Arguably, Seligman already recognized positive social relationships as contributors to happiness in his earlier work. Likewise, engagement in the sense of flow (i.e., being absorbed in the process of doing the work; Csikszentmihalyi, 1991) already implies accomplishment. Flow foretells success because the task is, by definition, tractable. In short, it is not clear how much the theory benefits from the expanded set of criteria.

Perhaps Seligman had other reasons for reworking his theory. In *Flourish* he writes that he "detest[s] the word *happiness,* which is so overused that it is almost meaningless. It is an unworkable term for science" (p. 10). The departure from the happiness model lies in the abandonment of the idea that there is a single underlying dimension along which individuals can be lined up. Here, Seligman says, he is breaking with Aristotle: "Authentic happiness theory comes dangerously close to Aristotle's monism because happiness is operationalized, or defined, by life satisfaction. Well-being has several contributing elements that take us safely away from monism" (p. 16). With the revision, Seligman is moving toward a family resemblance model, and a model that is not exclusively subjective. However, he says little about how outside observers or quantifiable criteria, such as measures of accomplishment (citation count) or the size of a person's social network (friends on Facebook), contribute to assessment. Hence, the break with Aristotle is more rhetorical than factual. If, as I suspect, positive social relationships and accomplishment respectively are features of positive affect and engagement, the old model is shaken but not broken.

In *Flourish,* Seligman proposes the notions of self-determination and free will as the new conceptual core. People who are well are people who "choose for their own sake" (p. 14) and thereby claim responsibility for their actions. Whereas there is plenty of research on self-control and intrinsic motivation, the notion of free will has no support in experimental psychology. It is an even less workable term for science than is the term *happiness* (Krueger, 2010; Miller, 2008).

The status of the empirical research supports the impression that there is no clean departure from *happiology* (Seligman's term). Assessment still depends on subjective, self-report instruments. Working with the U.S. Army, Seligman's colleagues have developed a Global Assessment Tool (GAT) to measure, via self-report, soldiers' "comprehensive fitness." The GAT taps into the domains of "emotional fitness," "social fitness," "spiritual fitness," and "family fitness." These domains are not well aligned with the five theoretical components of well-being, an issue that Seligman does not pursue. He also continues to promote use of the "signature strengths test," which predates his revisionist theory. The relationships between the 24 strengths and the five elements of the theoretical model remain unexamined.

The volume that reviews the character strengths is an interesting creation. Peterson and Seligman (2004) are listed as authors, whereas numerous other people are credited as "contributors." How much of the writing did these contributors do? The chapter on humility and modesty, for example, was written by Julie Exline and her colleagues, who then sold it to Peterson and Seligman, who then published it after minimal editing under their own names.

Seligman's revisionist project is incomplete in another way. The empirical evidence for his new model is just beginning to emerge. Much of this evidence comes from pilot-type efforts done in college classes. The book provides an index with references, but many of these are overview articles in the *American Psychologist,* coauthored by Seligman himself. Seligman is well aware that the gold standard of validation is true experiments or randomized trials done in independent labs. Unfortunately, there is little of that.

Throughout the book, Seligman offers a high dose of positive self-affirmation. He keeps reminding the reader (and thereby himself) of his academic credentials, such as his speedy completion of graduate school, or his run of federal funding, which he says has left him with more money than he knows how to spend. When recalling how his collaboration with the Army began, he portrays himself as the cautious scientist who wants to do more research on the training programs for psychological fitness before implementing them. It is here that he relates a revealing episode.

In 2008, at a "Seligman Lunch" at the Pentagon, he is told, "We have read your books, and we want to know what you suggest for the army" (p. 126). The chief of staff of the Army, "the legendary George Casey," announces that "Dr. Seligman here is the world's expert on resilience, and he's going to tell us how we are going to do it," that is, how "resilience will be taught and measured throughout the United States Army." Casey also says, "Dr. Seligman, Comprehensive Soldier Fitness began two months ago. It is under General Cornum's command" (p. 128). And,

Casey continues, "General Cornum, I want you and Marty [Marty!] to put your heads together, put flesh on the skeleton of Comprehensive Soldier Fitness, and report back to me in sixty days" (p. 129; brackets in the original). In short, Seligman has to scramble to catch up with the Army empirically, theoretically, and ideologically. He does nothing to defuse the impression that his revised theory of well-being is a response to what the Army had already chosen for him to do. Says Cornum, "If we had waited [for the science to catch up], we'd still be talking and planning" (Azar, 2011, p. 32).

Flourish is an inside account of Seligman's struggle to stay in control of a project he helped set in motion. He wants positive psychology to be his legacy, and he wants to plant one more footprint before the paradigm leaves Penn Harbor forever, as it eventually must. The appearance of the U.S. Army on the scene suggests that the ship may have already sailed. Consider this fact again: The U.S. Army controls the massive dataset on which Seligman stakes the future of psychology.

Flourish is also a very personal book. I suspect that there is more candor than the author intended. Throughout the book, Seligman makes his claim that he is a very, very important psychologist. If you missed the note on how he was elected president of the American Psychological Association by the widest margin of votes ever, he reminds you of it in the biographical blurb in the back. At the same time, he confesses to having self-doubts and fears of being a failure. Happiness theory has not worked for him, it seems.

This brings me to a structural problem of positive psychology. The public will probably ask any professed positive psychologist whether he or she personally represents the promise of the theory. This may be unfair, but it is hard to avoid. Practitioners of conventional (or "negative") psychology have it easier here. All they need to be is ordinary. Proponents of the positive face the risk of either falling short of their own ideals or, when trying too hard to prove that they do not fall short, appear messianic.

A related structural problem is that positive psychologists must compete with thinly credentialed pop psychologists who dominate the lucrative self-help book market. Conventional psychologists face no such competition and run no risk of being confused with popular motivation experts. Back in Akumal, Seligman announced that success has arrived when you have published your first book on self-help that really works. Judging from Seligman's

assertion that the crucial scientific questions will someday be answered by the Army data, it seems that positive psychology is not fully emancipated yet. But let us choose optimism!

Joachim I. Krueger
Department of Cognitive, Linguistic & Psychological Sciences
Brown University, Box 1853
89 Waterman St.
Providence, RI 02912
E-mail: Joachim_Krueger@Brown.edu

NOTES

1. The term *learned helplessness* may be a misnomer because it could be misconstrued as meaning that something useful has been learned. *Acquired helplessness* seems more apt.

2. Seligman does not discuss the question of whether his theory endorses metaphysical, supernatural, or religious beliefs as normative.

REFERENCES

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Azar, B. (2011). Positive psychology advances, with growing pains. *APA Monitor, 42,* 32.

Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience.* New York, NY: HarperCollins.

Gilbert, D. T. (2006). *Stumbling on happiness.* New York, NY: Knopf.

Krueger, J. I. (2010). A random walk through the free willderness. *Psychology Today.* Retrieved from http://www.psychologytoday.com/blog/one-among-many/201012/random-walk-through-the-free-will-derness

Miller, A. (2008). A critique of positive psychology—or "the new science of happiness." *Journal of Philosophy of Education, 42,* 591–608.

Overmier, J. B., & Leaf, R. C. (1965). Effects of discriminative Pavlovian fear conditioning upon previously or subsequently acquired avoidance responding. *Journal of Comparative and Physiological Psychology, 60,* 213–217.

Overmier, J. B., & Seligman, M. E. P. (1967). Effects of inescapable shock upon subsequent escape and avoidance responding. *Journal of Comparative and Physiological Psychology, 63,* 28–33.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification.* New York, NY: Oxford University Press.

Seligman, M. E. P. (1998). *Learned optimism.* New York, NY: Pocket Books.

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential and lasting fulfillment.* New York, NY: The Free Press.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist, 48,* 1117–1131.

Solomon, R. L., & Corbit, J. D. (1974). An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review, 81,* 119–145.

Solomon, R. L., & Wynne, L. C. (1953). Traumatic avoidance learning: Acquisition in normal dogs. *Psychological Monographs, 67*(4).

Weiner, B. (1974). *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Press.