

Rationality and Type 1 Processing

**Cognitive Unconscious and Human Rationality**

Edited by Laura Macchi, Maria Bagassi, and Riccardo Viale. MIT Press, 2016. 385 pp. Cloth, \$54.00.

The macro-themes and the micro-themes of this book elicit somewhat different reactions. The micro-themes—exemplified in the research programs described in several individual chapters—represent a stellar sample of the best cutting-edge research and theory in the science of human reasoning. These chapters, and I will touch on several, present incisive theoretical advances in their specific areas. Many discuss data from studies that make important theoretical distinctions. The reader of this volume will be exposed to a panoply of cutting-edge topics in reasoning research.

I am somewhat less comfortable with the general macro-themes that introduce the book and ostensibly tie together the multifarious chapters. This introduction of macro-themes is carried out in the five pieces that open the book: a Foreward by Keith Frankish; an introduction by the editors, Macchi, Bagassi, and Viale; and in three chapters that comprise a section of the book titled Bounded Rationality Updated (chapters by Gigerenzer; Elqayam; and Bagassi and Macchi). It is not so much that I disagree point-by-point with the authors of these stage-setting chapters, but that several reactions that they trigger, when taken collectively, make me a little ill-at-ease with how the volume is being presented from a macro-theoretical point of view. I will list my skeptical reactions first as categories and then elaborate on them.

My skeptical reaction #1 is that some of what the introductory authors present as a new, fresh consensus may well be a consensus, but it is not all that fresh—some of it having been around for 20 years or more. Skeptical reaction #2 is that many of these introductory chapters start out with bold claims, but as the chapter wears on, the claims are watered down to something much less bold—indeed into something that is pretty widely assumed. Often, the presentation of the older theoretical views that are being superseded are somewhat caricatured. That strategy certainly makes the new views seem more correct, but as soon as one realizes what a caricature the views they superseded are, then one becomes less impressed. Reaction #3 is that there is a lot of uncertainty about when authors are talking about the descriptive, the prescriptive, or the normative. Reaction #4 is that when evolutionary and adaptationist arguments come into play, there is often a slipperiness in the discussion about whether we are talking about the subpersonal level of the genes or the level of the whole organism (a point I will come back to; see Stanovich 2004).

Frankish sets an exciting context for the volume in his Foreward. However, the very same Foreward that excites me, sets off a few of the skeptical reactions that I have just listed. Frankish sets forth some of the positive developments that he sees as deriving from the new perspective on reasoning research represented by this book. For example, he states that “these developments do undermine the simple equation of implicit processing with bias and explicit processing with normative correctness” (p. ix). But it has repeatedly been pointed out that the equation of Type 1 processes with all bad thinking and Type 2 processes with correct responding is the most persistent fallacy in the history of dual process theory (now reaching its fortieth anniversary; Posner & Snyder, 1975;

Shiffrin & Schneider, 1977; Wason & Evans, 1975). Evans and Stanovich (2013) pointed to some early theoretical treatments (e.g., Epstein, 1994) as incorrectly labeling their two processing modes to suggest this fallacy. Indeed, the early originators of the heuristics and biases research tradition—the tradition that did much to fuel the popularity of dual-process theories—consistently reiterated that Type 1 processing modes often lead to normative responding and efficient task performance (Kahneman, 2000, 2011). So while Frankish is right to point out this implication of the reasoning perspective contained in this volume, the implication itself is not new.

Likewise, Frankish argues that the perspective represented by these chapters views “the heuristics we intuitively rely on as not some optimal compromises but often the best solutions to particular context specific problems” (p. ix). This is correct, but it has been part of the general reconciliation between the heuristics and biases tradition and the evolutionary/adaptation tradition of interpreting reasoning errors for some time now. This reconciliation and synthesis is present in works well into their second decade now, including, in chronological order: Stanovich (1999 2004), Stanovich and West (2000), Kahneman and Frederick (2002), and Samuels and Stich (2004).

The synthesis follows from interpreting the responses primed by Type 1 and Type 2 processing as reflecting conflicts between two different types of optimization—fitness maximization at the subpersonal genetic level and utility maximization at the personal level. The synthesis acknowledges that the evolutionary psychologists have often shown that the adaptive response on a particular task is the modal response on the task—the one that most subjects give. However, that data pattern must be reconciled with another

finding often obtained: that lower cognitive ability is often associated with the response deemed adaptive on an evolutionary analysis (Stanovich, 1999; Stanovich & West, 2000; Stanovich et al., 2016). The synthesis argues that the evolutionary interpretations do not impeach the position of the heuristics and biases researchers that the alternative response given by the minority of (more cognitively able) subjects is rational at the level of the individual. Subjects of higher analytic intelligence are simply more prone to override Type 1 processing in order to produce responses that are epistemically and instrumentally rational. This rapprochement between the two camps was introduced by Stanovich (1999).

As we proceed through the Foreword by Frankish, we see more conclusions that are right in line with the synthesis. For example, when Frankish says

“the better we understand the functions of the human reasoning system and its power and limitations, the better placed we shall be to tweak it to do the more artificial jobs we sometimes require of it. If you must cook an egg on a toaster that helps to understand exactly how a toaster works” (p. ix). This is precisely what the old synthesis going back to Stanovich (1999) argued. That in the hostile environments of modernity we often have to cook an egg on a toaster and this is why over 40 years ago Tversky and Kahneman became interested in errors in hostile environments (Lewis, 2017; Stanovich, 2004).

The Introduction by Macchi, Bagassi, and Viale, continues the trend of promising newness but actually delivering a solid consensus view—and I do not mean to be critical of the consensus view that is delivered. It is indeed theoretically and empirically justified. It is just that it is more in line with other approaches than is implied in many of the book’s chapters. The

book's title meshes the idea of the cognitive unconscious with the issue of rationality. Minus the conscious/unconscious distinction, many of the chapter contributions amount to showing that Type 1 processes are efficacious for the purposes for which they were designed and often normatively correct as well. That point is well made here and developed in glorious detail in many of the chapters, but the point is an elaboration and continuation of a stance going all the way back to the original Kahneman and Tversky papers, not a new insight.

In the Introduction, Macchi, Bagassi, and Viale take us on a very useful tour of the many kinds of rationalities mentioned by the various authors, including: ecological rationality, grounded rationality, bounded rationality, interactional rationality, and others. These are too many rationalities to be helpful, and they interact to lead to the confusion that I mentioned above about what level of analysis we are talking about; for example, whether the discussion is about a subpersonal process or about the personal level of analysis. Definitions of rationality must be kept consistent with the entity whose optimization is at issue.

Another reason not to proliferate the number of rationalities we are discussing is to simply look at the sorry history of intelligence research which tended to do the same thing—that is, to attach the term intelligence to a variety of other cognitive entities (Stanovich, 2009). Major theorists such as Sternberg and Gardner talk about practical intelligence, creative intelligence, interpersonal intelligence, bodily-kinesthetic intelligence, etc. In such usages, the word intelligence becomes a marker for “optimal or expert behavior in some domain”. So, for instance, when Sternberg discusses high practical intelligence it can be translated to mean “optimal behavior in the domain of practical affairs” or when Gardner talks about high bodily-

kinesthetic intelligence he means little more than high functioning in the bodily-kinesthetic domain. The word intelligence is actually superfluous. It is there merely to add status to the domain in question. Ironically, calling everything intelligence drains the term intelligence of meaning (Stanovich, 2009). The proliferation of these different “rationalities” worries me in the same way.

Gigerenzer’s chapter leads off a meta-theoretical opening section titled Bounded Rationality Updated. His chapter gives us a useful exegesis of Herb Simon’s use of the term bounded rationality and describes how the term is sometimes misused (or at least variously used) in cognitive science and in the heuristics and biases literature in particular. It is useful to have the bounded rationality term situated and to realize that it has been overused or not sufficiently clarified by many theorists. However, the necessary clarifications do nothing to undermine the contributions of the heuristics and biases literature. Nor do they impeach anything in dual-process theory or any other foundational view in the current loose consensus in reasoning research.

Perhaps Elqayam’s chapter in this section is meant to be read in a more revolutionary way, but it again has the property that I alluded to above. The chapter starts out with somewhat bold claims that end up getting watered down as the chapter reaches its conclusion. Some of the early assertions come close to caricatures. No one in cognitive science has viewed rationality as entirely encompassed by logic for some time now. Like Gigerenzer’s critique of the misuse of the bounded rationality term, Elqayam’s critique of evaluative language in psychology does nothing to endanger the contributions of research in the heuristics and biases tradition (Stanovich, 2004).

The sharpness of the attack on normative models is greatly reduced at the end of the chapter where Elqayam introduces her terminology of grounded rationality. The core idea here is that rationality depends on the agent's goals and the context in which the agent operates. I have absolutely no objection to this, because it is completely consistent with the notions of rationality that are widespread in cognitive science (Stanovich, 2012). The grounded rationality notion greatly waters down the critique of normativity that precedes it in the essay, bringing the view in this chapter more in line with the 20-year-old consensus that I referred to above (Stanovich & West, 2000). At the end of the chapter, Elqayam comes close to endorsing a two-tiered rationality assessment of the type I have argued for previously (Stanovich 2013)—one that includes a broad meta-critique of the desires pursued as well as an assessment of the narrower instrumental rationality that is achieved.

To round out the three chapters providing a meta-theoretical context for the book, Bagassi and Macchi argue for a variety of features that should be emphasized in the study of rationality. These include giving an implicit thought a more central role in our theories and emphasizing the automatic contextualizing features of Type 1 processing. They argue that moves in this direction would give this research area more originality and they refer to much interesting work on the Type 1 processes involved in solving insight problems. The chapter also argues for the importance of the interpretive function in reasoning and that it is an adaptive characteristic of human thought (this is a theme that is taken up in even more detail in the chapter by Mercier, Bonnier, and Trouche). They refer to Levinson's (1995) emphasis on interactional intelligence.

In my very first stab at constructing a version of a dual-process model (Stanovich, 1999) I also relied heavily on Levison's (1995) notion of interactional intelligence, and I demonstrated how it played a role in setting up the conflicts that subjects experience on heuristics and biases tasks. The tendency to see designed conversational features in these problems was one of many so-called fundamental computational biases that I built into that early dual-process model. I discussed how Levinson (1995) and Adler (1984, 1991) had already laid out the fundamental continuum of abstraction assumptions versus interactional assumptions that explained how subjects approached these tasks. My own work invoked individual difference measurement to try to predict the likely default assumptions of different groups of subjects. This is all by way of noting that the points that Bagassi and Macchi emphasize are indeed valid, but many have been part of the general consensus for some years now.

The chapter by Mercier, Bonnier, and Trouche continues the emphasis on interactional intelligence as a suite of Type 1 processes that underlie observed reasoning behavior. They discuss a theory of reasoning that is grounded in the logic of the evolution of communication. The theory is subtle and nuanced. It explains things like the asymmetries in argument evaluation that are at the heart of most of myside bias paradigms. They discuss argument failures that are due to the curse of knowledge, or that arise because of egocentrism. Again, it is not that these discussions are wrong in any way, it is just that the authors need to emphasize the incremental nature of these contributions. They are evolutionary rather than revolutionary.

There is a long history of related views. Nozick (1993) has argued that in prehistory, when mechanisms for revealing what is true about the

world were few, a crude route to reliable knowledge might have been just to demand reasons for assertions by conspecifics. Dennett (1996, pp. 126-127) has argued that the need to respond to the justification queries of other conspecifics helped shape the internal mental inquiry processes in the direction of discrete categories—categories that mesh well with the properties of language-based thought. These discrete categories fostered by language then become an ideal medium for representing one's own thoughts to oneself (see Carruthers, 2006). Sterelny (2001) develops similar proposals in attempting to argue that social intelligence was the basis of early simulation ability to the social intelligence hypothesis (see also, Mithen, 1996, 2000; Nichols & Stich, 2003).

Some years ago, philosopher Allan Gibbard (1990) anticipated all of these proposals with his emphasis on consistency checking among conspecifics—“the need to work through normative commitments in community” (p. 74)—and his view that “to prepare oneself to meet demands for consistency may require a strong imaginative life. A person will engage in imaginative rehearsal for actual normative discussion; he practices by holding himself to consistency” (p. 74). All of these views are, despite subtle differences between them, sketching the gene/cultural co-evolutionary history (Richerson & Boyd, 2005) of the negotiation of argument with conspecifics. The grounding of some of our reasoning biases in the logic of the evolution of communication goes back at least to Humphrey (1976).

Mercier, Bonnier, and Trouche emphasize perspective-taking within their view—that the speaker often fails to anticipate that the interlocutor does not hold a given **belief**—and that this is behind phenomena such as the curse of knowledge. This type of egocentrism as a miserly default falls

Dom Massaro 4/27/2017 5:08 PM

Deleted: believe

within the purview of theorizing on focal bias (serial analytic cognition that relies on a single focal model that triggers all subsequent thought, see Stanovich, 2011) that is a recurring theme in cognitive science. The notion of a focal bias conjoins several closely related ideas in the literature: Evans, Over and Handley's (2003) singularity principle, Johnson-Laird's (1999, 2006) principle of truth, focusing (Legrenzi, Girotto, & Johnson-Laird, 1993), the effect/effort issues discussed by Sperber, Cara, and Girotto (1995), and finally, the focalism (Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000) and belief acceptance (Gilbert, 1991) issues that have been prominent in the social psychological literature. Focal bias conjoins many of these ideas under the overarching theme that they all have in common—that the information processor deals with the most easily constructed cognitive model and is loathe to construct alternative models.

Mercier, Bonnier, and Trouche set their findings within an Apologist context (see Stanovich 1999) when they call such processing defaults as myside bias an “apparent” limitation of the reasoning mind (p. 205). This relates to a point that Kahneman and Tversky (1996) made in response to Apologist and Panglossian critics many years ago who were claiming they could make certain cognitive illusions “disappear”. To show that processing errors that occur without a supportive context can be reduced when supportive cues are introduced does not make the bias any less an error when it occurs in the absence of supportive cues. Likewise, to show that a bias occurs for sound evolutionary reasons does not mean that it is any less an error when it shows up in a hostile modern environment that requires that it be overridden (Stanovich, 2004).

The point (extensively discussed by Kahneman, 2011) is that an attribute-substituting System 1 and a lazy System 2 can combine to yield

rational behavior in benign environments but can yield seriously suboptimal behavior in hostile environments. A benign environment is an environment that contains useful cues that, via practice, have been well represented in System 1. Additionally, for an environment to be classified as benign, it must not contain other individuals who will adjust their behavior to exploit those relying only on System 1 heuristics. In contrast, a hostile environment for heuristics is one in which there are no cues that are usable by System 1 (causing the substitution of an attribute only weakly correlated with the true target). Another way that an environment can turn hostile is if other agents discern the simple cues that are triggering the cognitive miser's System 1—and the other agents start to arrange the cues for their own advantage (for example, advertisements, or the deliberate design of supermarket floorspace to maximize revenue).

None of my critique of some of the metatheory in the volume is meant to deny the contribution of the compelling research and theory presented in many of the chapters. Several chapters, such as those by Over and by Baratgin and Politzer, discuss the new theoretical paradigms in logical reasoning, particularly reasoning about conditionals and the interpretation of that reasoning within the context of probabilistic models. Several chapters report seminal contributions to the study of creativity (chapters by Bagassi & Macchi; Gilhooly; and Viale) and the integration of information during decisions (Betsch, Ritter, Lang, & Lindow). Hilton's chapter drills down deeply into the pragmatics of conditionals, and Sun's chapter helps to clarify part of what we mean by reflective and intuitive by presenting a dual-process architecture that is different from the more common default interventionist approaches in the dual-process literature.

Bonnefon and Billaut provide a review of the Great Rationality Debate in dual-process theory (Evans & Stanovich, 2013) in the context of individual differences. They argue for the study of individual differences that are beyond the traditional cognitive abilities and thinking dispositions that have been studied to date. They report intriguing investigations of several other personality traits and a quite focused investigation of just-world belief. Sloman and Barbey make a case for a foundational view of human intuition that is based on the assumptions that the world is law governed, that it is deterministic, and that free will is real. Oaksford's chapter contains an elegant discussion of another assumption that I would take as part of the consensus view: that in the real world, there is a trade-off between achieving our goals in a timely manner and in the search for truth. He explains the logic of how explicit verbal reasoning intervenes between perception and action in practical contexts when goal fulfillment is in jeopardy. This leads him to a view of human cognition where we achieve practical rationality in most everyday contacts but we have to stray from epistemic rationality along the way. There is a final chapter of Closing Thoughts that consists of writings by Giuseppe Mosconi (to whom the volume is dedicated) that are of relevance to many of the themes of the volume.

The chapter by Ball and Stubble presents a detailed analysis of work on the belief bias effect in syllogistic reasoning. They argue that extant data are best fit by what they claim is a new hybrid dual-process model advocated by De Neys. The model seems fine to me, but it can only be considered new if we grant that one of its assumptions has never been made before. That assumption is that Type 1 processing includes what De Neys calls logical intuitions, in addition to intuitive heuristics. However, I do

not think that positing so-called logical intuitions is a new addition to dual-process theory at all. It has been there all the time, but has simply been called by different terms. Almost forty years ago, Shiffrin and Schneider (1977) talked about compiled control processes becoming automatic. In my own theorizing, I have often acknowledged the existence of highly compiled mindware in System 1 that supports the normative response (Stanovich, 2009). As early as 2000, in response to the commentators on our *Behavioral and Brain Sciences* target article, West and I noted that “some of the well framed intuitions referred to by Teigen may well be acquired intuitions—having their origins in capacity-intensive serial processing, yet now having the encapsulated, automatic characteristics of modular processes. Some statistics instructors, for example, become unable to empathize with their students for whom the basic probability axioms are not transparent. The instructor can no longer remember when these axioms were not primary intuitions” (pp. 709-710, Stanovich & West, 2000).

In short, the hybrid dual-process model discussed in this chapter is only new if we make the assumption that all previous dual-process models carry the assumption that only non-normative biases are stored in System 1 and never the normative response. As Evans and Stanovich (2013) point out, dual-process theorists have been at pains to correct this misconception for many years now. Depending on the task and person, normative mindware is more or less instantiated in System 1 on a continuum (Figure 3.2 in Stanovich, West, and Toplak (2016) represents the latest dual-process thinking on this issue). The continuum ranges from situations where no mindware has been instantiated in System 1 or System 2, to situations where the normative response is triggered automatically from System 1 because the normative response is so well instantiated there.

The continuum also demarcates the area explored in De Neys (2014) work—where normative mindware, partially instantiated in System 1, competes with the intuitive heuristics in System 1. The hybrid dual-process theory is only new if previous dual-process theories are conceived as strawmen where no normative information can ever enter System 1.

These quibbles aside, Ball and Stupple present a very viable model to explain the belief bias findings. The hybrid model with multiple and conflicting outputs coming from System 1 obviously needs processes of regulation and control. Ball and Stupple rightly point to work on meta-reasoning as relevant to this part of the model and specifically direct us to the work of Valerie Thompson on meta-control and regulation in situations of cognitive conflict (whether within the context of a dual-process model or some other kind of architecture; regulation mechanisms are needed either way). Fortunately for the reader, there is a chapter by Thompson herself, along with collaborators Nicole Therriault and Ian Newman included in the volume, and this chapter alone is worth the price of the book. The chapter more than achieves its goal of providing a comprehensive but also nuanced review of work on meta-reasoning, including what is known about both monitoring and control processes. Their review starts us at exactly the right place, theoretically: trying to understand when reasoners are sometimes content with the default answer and when they are not. Such an understanding involves individual difference issues of the person as well as issues of the task situation. The authors take us through important literatures on judgments of solvability, judgments of ongoing confidence, feeling of rightness judgments, and final judgments of confidence. One recurring point in their review is that Type 2 processing is often engaged in response to implicit cues. This point makes the chapter by Thompson,

Therriault, and Newman one of the most integrative of the volume in terms of tying together the concepts of the book title: the cognitive unconscious and human rationality.

A particularly important insight from the concluding sections of this chapter concerns how we view autonomous processes (the key feature of Type 1 processing; see Stanovich 2011; Evans & Stanovich, 2013). Thompson and colleagues point out that Type 2 processing might share at least one aspect of Type 1 processing; namely, that of being *initiated* autonomously. The points of diversion between Type 1 and Type 2 processing are not at the *initiation* stage. The differentiation between these types of processing concerns whether or not the processing is ballistic; that is, whether it must run to completion without the possibility of interruption (Type 1 processing) or whether the reasoner may alter the length of time that the process runs, including the possibility of terminating it (Type 2 processing).

Keith E. Stanovich

Department Applied Psychology and Human Development

University of Toronto

252 Bloor St. West

Toronto, Ontario

Canada M5S 1V6

Email: keith.stanovich@utoronto.ca

#### References

Adler, J. E. (1984). Abstraction is uncooperative. *Journal for the Theory of Social Behaviour*, 14, 165-181.

- Adler, J. E. (1991). An optimist's pessimism: Conversation and conjunctions. In E. Eells & T. Maruszewski (Eds.), *Probability and rationality: Studies on L. Jonathan Cohen's philosophy of science* (pp. 251-282). Amsterdam: Editions Rodopi.
- Carruthers, P. (2006). *The architecture of the mind*. New York: Oxford University Press.
- De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning, 20*, 169-187.
- Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. New York: Basic Books.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709-724.
- Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2003). A theory of hypothetical thinking. In D. Hardman & L. Maachi (Eds.), *Thinking: Psychological perspectives on reasoning* (pp. 3-22). Chicester: Wiley.
- Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223-241.
- Gibbard, A. (1990). *Wise choices, apt feelings: A theory of normative judgment*. Cambridge, MA: Harvard University Press.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*, 107-119.
- Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303-317). London: Faber & Faber.
- Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology, 50*, 109-135.

- Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press.
- Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. *Behavioral and Brain Sciences*, 23, 681-683.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus & Giroux.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582-591.
- Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (1993). Focussing in reasoning and decision making. *Cognition*, 49, 37-66.
- Levinson, S. C. (1995). Interactional biases in human thinking. In E. Goody (Eds.), *Social intelligence and interaction* (pp. 221-260). Cambridge: Cambridge University Press.
- Lewis, M. (2017). *The undoing project*. New York: Norton.
- Mithen, S. (1996). *The prehistory of mind: The cognitive origins of art and science*. London: Thames and Hudson.
- Mithen, S. (2000). Palaeoanthropological perspectives on the theory of mind. In S. Baron-Cohen, H. Tager-Flusberg (Miyake *et al.*, 2001)g, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd Ed.)* (pp. 488-502). Oxford: Oxford University Press.

- Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Oxford University Press.
- Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). New York: Wiley.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Samuels, R., & Stich, S. P. (2004). Rationality and psychology. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 279-300). Oxford: Oxford University Press.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.
- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *57*, 31-95.
- 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: 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. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.

- Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. Holyoak & R. Morrison (Eds.) (pp. 343-365), *The Oxford handbook of thinking and reasoning*. New York: Oxford University Press.
- Stanovich, K. E. (2013). Why humans are (sometimes) less rational than other animals: Cognitive complexity and the axioms of rational choice. *Thinking & Reasoning*, 19, 1-26.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-726.
- Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The rationality quotient: Toward a test of rational thinking*. Cambridge, MA: MIT Press.
- Sterelny, K. (2001). *The evolution of agency and other essays*. Cambridge: Cambridge University Press.
- Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition*, 3, 141-154.
- Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78, 821-836.