CHAPTER

1

Rationality and its Contexts

T.J. Lane

Taipei Medical University, Graduate Institute of Humanities in Medicine, Taipei, Taiwan; Taipei Medical University–Shuang Ho Hospital, Brain and Consciousness Research Center, New Taipei City, Taiwan; Academia Sinica, Institute of European and American Studies, Taipei, Taiwan; National Chengchi University, Research Center for Mind, Brain and Learning, Taipei, Taiwan

In the centuries-long cultural project of humanity's growing awareness of its place in the universe, evaluating the extent of human rationality represents a seminal project. ***Keith E. Stanovich,*** **The Robot's Rebellion, *2005, 92.***

A cursory glance at the list of Nobel laureates for economics is sufficient to confirm Stanovich's description of the project to evaluate human rationality as seminal. Herbert Simon, Reinhard Selten, John Nash, Daniel Kahneman, and others were awarded their prizes less for their work in economics, per se, than for their work on rationality as such. Although philosophical works have for millennia attempted to describe, explicate, and evaluate individual and collective aspects of rationality, new impetus was brought to this endeavor over the last century as mathematical logic along with the social and behavioral sciences emerged. Yet more recently, over the last several decades, propelled by the emergence of artificial intelligence, cognitive science, evolutionary psychology, neuropsychology, and related fields, even more sophisticated approaches to the study of rationality have emerged.

Some among these new lines of inquiry, including those pursued by Kahneman (Kahneman & Tversky, 1972), have been interpreted as implying that humans are not quite the rational creatures that Aristotle imagined us to be (Piattelli-Palmarini, 1994). Indeed, it does seem to be the case that our performance in rational choice experiments is suboptimal, falling short of expected utility or Bayesian prescriptions; sometimes even experts—medical, legal, or engineering—perform at surprisingly and

Rationality. http://dx.doi.org/10.1016/B978-0-12-804600-5.00001-5

alarmingly poor levels on tests directly relevant to their areas of expertise (Kahneman & Tversky, 2000; Nisbett & Ross, 1980; Kahneman, Slovic, & Tversky, 1982). To cite just one example, trained physicians regularly and unnecessarily scare the bejeebers out of patients by committing the base rate fallacy: that is, they exaggerate the significance of positive results in diagnostic tests for relatively rare medical conditions (Hamm, 1996). But what these suboptimal performances imply about the nature of rationality more generally is contentious, having triggered what is in some quarters referred to as the "rationality wars"(Samuels, Stich, & Bishop, 2002).

It might seem that waging academic "war" over scientific findings, even if they do have important implications for how we should regard the consultations of experts, is excessive. But more is at stake than assessment of expertise. Since the assumption that humans are rational is pivotal to our understanding of the type of creature that we are, any significant challenge to that assumption might be felt demoralizing, not unlike the realizations that our species and solar system are products of evolution. Implications of these findings are, potentially, sweeping. One worry is that rationality has been thought by some to be a precondition for many of the capacities that humans exhibit (Davidson, 1984). A second, perhaps more troubling, worry is that if it is the case that we fall very far short of Aristotle's notion of rationality, such a finding might necessitate a dark reassessment of the human capacity for moral responsibility (Dahan-Katz, 2013). There was a time when concerns about moral responsibility were motivated more by the place of humans in the causal structure of the world. Recently, however, the focus of much attention has shifted to rationality and the psychological capacities that constitute its foundation (Morse, 2007), those very capacities that have been called into question. So getting clear about how to interpret the relevant scientific findings is no small thing.

We should be clear, however, about what is not contentious: substantial deviations from subjectively expected utility and Bayesian models are commonplace (Stanovich, 2010). We are not as responsive to reason when choosing beliefs or actions as may once have been thought, or at least hoped. At the personal, conscious level, we are not proficient implementers of Bayesian inference, even though it seems our brains (the subpersonal, nonconscious level) often act like a Bayesian mechanism (Hohwy, 2013). That we are susceptible to personal-level, systematic cognitive bias is no longer subject to dispute. There are limits or "bounds" to human rationality: in the context of decision-making, individual or organizational, what might be optimal is not necessarily what should be expected, for there are always constraints on information-processing capabilities (Simon, 1972, 1983). This is an instance wherein the ought-implies-can principle must be invoked: how we should reason is necessarily constrained by the capabilities of brains and their environments, no matter whether the brains are artificial, nonhuman, human, healthy or

unhealthy, and irrespective of whether we are practicing science or reflecting on the seeming irrationality of recondite philosophical texts.

Many among these constraints are temporal, constraints that are most tellingly evident when agents must decide whether or how to act (practical rationality), since such decisions are made in real time. Indeed, even polynomial time greatly constrains the capacity for evaluating belief, despite the fact that within complexity theory polynomial expressions are relatively easy (Ladner, 1975). Other constraints involve the dynamic adjustment of our aspirations to features of unfolding context; decision alternatives for what to believe or how to act are not fixed or predetermined, in the way that seems to have been implied by the von Neumann–Morgenstein utility theorem or other similar normative models that characterize human beings as utility maximizers (von Neuman & Morgenstern, 1944). Instead, our aspirations are discovered in the process of searching. These, then, are adjusted upward or down, in a manner befitting the context of discovery (Selten, 2001).

Obviously, search cannot be unbounded, at least not literally. Normative models like that proposed by von Neumann and Morgenstern posit exhaustive coverage of alternatives, but neither mere mortals nor mere machines can conduct exhaustive searches, especially when it is not even clear just what might count as an alternative. Toward this end, to prune the tree of possibilities and reduce the burdens of evaluating each, we rely heavily upon rules of thumb or heuristic strategies (Gigerenzer & Goldstein, 1996). Often too, the framing of a problem (Tversky & Kahneman, 1981), conversational context (Kahneman & Tversky, 1982, pp. 132–135), and emotion (Damasio, 1999) efficiently, though not necessarily reliably, reduce search space.

Even if there is consensus on goals and their evaluation, search is not made significantly less difficult. In other words, even if we presuppose that the utilities of goals have been antecedently fixed, the problems confronted by mortals and machines do not disappear (Hempel, 1965, pp. 463–472). Deciding how to act, for example, will vary among persons, reflecting different inductive attitudes or degrees of optimism. Mathematical models of decision-making under uncertainty, like the "maximin" and "maximax" rules, reflect these differences. The former represents extreme caution, in that the maxim for action is "assume the worst possible outcome"; the latter, optimism, in that the maxim for action is expect "the best possible outcome."

Simon (1996, pp. 27–30, 119–121) neatly encapsulates many of these ideas with the felicitous, if not altogether aesthetically pleasing, term "satisfice." Rather than optimizing or maximizing, humans tend to satisfice, that is, we tend to choose between better or worse. Choosing that which is best is usually not an option because we rarely have a method of finding the optimum. Bound by practical computational limits, we cannot

generate the entire list of possible alternatives nor fully evaluate the merits of those we do manage to generate. These difficulties are compounded by the fact that even were we to stumble upon the best alternative early in the search process, we could not recognize it as the best until after we had generated the entire list. So we satisfice; we conduct moderate, not exhaustive, searches until we find something acceptable, not optimal.

Naturally enough, however, the capacity to search effectively varies greatly. We differ in temperament and talent, as well as in the particular enculturation practices to which we are exposed. These differences matter although in certain important respects, we all must settle for satisficing, people exhibit better or worse capabilities for thinking up new and fruitful possibilities. Like Selten and Simon, Nozick (1993, pp. 172–181) observes that not restricting our search for alternatives to a preexisting set can be a good strategy. Even an inferior choice from among a newly generated list might be better than the best choice from a preexisting list. That which Nozick dubs "rationality's imagination" is important because lacking imagination our searches would not only fail to be exhaustive, they would be myopic. Not all who try to imagine new alternatives will succeed; more than likely a majority will fail. But rationality is as much social as it is individual (Weber, 1978; Lazarsfeld & Oberschall, 1965; Clark, 2001), and one important implication of this fact is that the costs of individual failure do not necessarily affect the group. On the other hand, successes achieved by individuals can have positive effects for the group. Thus society's capacity for insulating the group from individual failure contributes directly to promotion of rationality's more exploratory and risk-incurring features.

It should be acknowledged that von Neuman–Morgenstein utility and similar models have a well-deserved workhorse status, especially within modern economics (Karni, 2014). Nevertheless, it is clear that one among their limitations is the requirement "to abstract away aspects in the contextual environment" (Stanovich, 2005, p. 247). But the constraints that result in our suboptimal, satisficing performance make it abundantly clear that context does matter. Botterill & Carruthers (1999, p. 107) describe the implications of recent empirical and conceptual research thus: "…standards of rationality for belief-forming processes should be relativized to our needs as situated, finite, enquirers after truth." Indeed, a principal goal of this volume is to promote and assist with the task of relativizing the belief-forming processes indicative of rationality to context, features of the environment that utility maximizing models disregard.

It might be objected that enquiring "after truth," sets a goal no less impractical, or conceptually misguided, than exhaustively searching among a tree of possible alternatives. But even scholars who are wary of setting off in search of truth, like Thomas Kuhn, invoke "rationality" and tout the virtues of *rational* theory choice (Kuhn, 2000a). Kuhn is dubious about existing theories of rationality (Kuhn, 2000b, p. 159), viewing them as in

need of "readjustment." But he takes explaining the success of science with respect to enhanced efficiency in puzzle solving as a possible starting point for developing new theories of rationality. A principal reason for the organization of this volume is the editors' belief that even if we assume that puzzle solving, or some similar version of "temperate" rationality (Newton-Smith, 1981, pp. 266–273), is a proper platform upon which to begin the task of "readjusting" theories of rationality, puzzle solving itself would still need to be embedded in constraint-sensitive contexts.

Of course Kuhn's discussions of rationality are confined to the context of scientific practice. Given the premise that belief-forming processes need to be contextualized, perhaps it would be better to abandon the project of investigating rationality, in general, and replace it by approaches distinctive of individual disciplines. But to take such a step would be imprudent because even though in many discussions of rationality the normative-descriptive distinction is not sharply drawn (Richardson, 1998, p. 567), there is no basis for treating "rationality" as a purely descriptive, natural kind term like, say, "fire," which underwent radical reclassification after the discovery that burning wood, rusting iron, and biological metabolism all involve oxidation, whereas the sun, lightning, northern lights, and fireflies do not (Churchland, 2002, pp. 129–131). There is no reason to anticipate discovery of an analogue to oxidation; ipso facto, there is no reason to expect that a similar compartmentalized reduction of "rationality" is in the offing.

But warning that investigations of "rationality" should not be segregated is grounded in more than an intuition that its conceptual status is unlike "fire." Although "rationality" remains somewhat diffuse, we have good reason to believe that it is a disposition shared by all cognitive agents, an important factor in marking the distinction between, say, rendering judgments and being lost in reverie (Byron, 2004). To say that rationality plays more of a role in the former than it does in the latter is to say something substantial. To a first approximation, it is that in virtue of which agents adopt or act upon beliefs, appropriately. Despite the obvious imprecision of this definition, it seems to identify a perfectly general capacity for forming true beliefs and performing successful actions, a capacity that transcends the presuppositions of any particular context or community, a capacity that is applicable to individuals or groups, to political economists or scientists, to peoples of all sociohistorical contexts, perhaps even to nonhumans (Trigg, 1993, p. 62). What is more, it is that which is lost or diminished when agents suffer from certain pathological conditions; in fact, it is often a symptom of those conditions (Bortolotti, 2010).

This general capacity seems to be what Quine (1976, p. 233) is referring to when he writes "science is itself a continuation of common sense. The scientist is indistinguishable from the common man in his sense of evidence, except that the scientist is more careful." Scientists are dependent

upon a "primitive sense of evidence," that they use "carefully and systematically." Both the common man and the scientist are agents who adopt or act upon beliefs, appropriately, albeit while exhibiting contextually shaped standards for what counts as appropriate. The "primitive sense of evidence" is a reflection of our capacity for rationality; "care and system" are what enable us to overcome—in a manner that is relativized to context—the many cognitive biases that have recently been described. It is that same care and system that make possible a bootstrapping of ourselves from common sense rationality to scientific rationality.

Quine (1976, p. 234) proceeds to observe that even the very preference for simplicity "is a lay habit carried over by science." He adds that the simpler of two hypotheses is generally regarded "not only as the more desirable but also as the more probable" (1976, p. 255). The latter point, concerning the greater probability of a simpler hypothesis being true, appears to be a point at which Simon and Quine converge, satisfice and simplify find common ground. In Simon's terms, we satisfice rather than optimize, but seem none too much the worse for having done so. Perhaps our bounded rationality is of a piece with our preference for simplicity, and it is a reliable guide to navigating this uncertain world because we have developed sufficiently simple and successful strategies. These strategies are foundational, part of the common core that makes an interdisciplinary approach to rationality so apt.

Nevertheless, there is no denying that the attempt to understand rationality in a way that is not bound to any one domain or tradition is partially influenced by philosophy's predilection for an expansive approach to scholarship. In the words of Sellars (1962, p. 1), philosophy aims "to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term." Although many of the essays collected here are animated by empirical work, this is primarily a philosophical work, one that aims to contribute to understanding how scientific, pathological, nonhuman, pedestrian, and other forms of rationality, even serious meditations on irrationality, "hang together."

To place this volume in historical context, it is an attempt to help make sense of an entire, millennia-long enterprise that has undergone an abrupt, somewhat rude awakening in recent decades. To a considerable extent, economists and philosophers have aligned with one another in a tug-of-war with psychologists over "rationality." Much but not all of the tension is more apparent than real because philosophers and economists have tended to emphasize theoretical and normative aspects (Rescher, 1988), while psychologists have tended to emphasize its practical and descriptive aspects (Kahneman, 2011). Nevertheless, it would be foolish to disregard the impact of recent psychological studies; after all, their having garnered a lion's share of Nobel Prizes for economics is not without warrant. But the evident influence of those studies does

not justify wholesale refutation of positions adopted by a generation of economists and philosophers.

The proper response to conflicts over "rationality" is, instead, to call in a plumber. Midgley (2001, p. 37) opines that "when the conflicts get so bad that we do notice them, we need to call in a philosophic plumber." The rationality wars were, in their more heated, overwrought moments, just that bad. In those moments, amid the handwringing over potentially "bleak implications for human rationality" (Nisbett & Borgida, 1975), a niche was created, a need for philosophers to plumb. Now that much of the dust has settled, toward what ends should the philosophic plumbers direct their efforts?

Acknowledging that belief-forming processes are constrained does not imply that constraints cannot be satisfied locally, or in a manner that is relativized to specific contexts (Glymour, 1992, pp. 361–363). We can apply what has been learned in recent decades to specific contexts, searching for, plumbing, perhaps even correcting some erroneous beliefs and errors in belief formation. Both in spite of and because of our bounded rationality this is doable. Since search space is restricted we know where to look, and because of subjective expected utility and Bayesian models, as well as, our familiarity with cognitive bias, we know what to look for.

Although the focus of this volume is on rationality as it is realized, or fails to be realized, in specific contexts, we do as well keep an eye on how distinct pockets of localized rationality "hang together," reflecting rationality "in the broadest possible sense of the term." After all, global considerations are often called upon to override local; what is best or better provincially and provisionally might not be abiding or perdurable in the broader scheme of things. Central banks and citizens might believe printing money a good way to revive depressed economies, until inflation sets in, negating the benefits of having more currency in circulation and triggering fears among neighbors of sovereign default.

Even apparent pathologies of rationality can be informative as regards how to comprehend rationality in the broad sense of the term. It is often the case that when normal functions are damaged the nature of remaining capabilities can be seen with greater clarity (Frith, 1998). To cite just two examples, investigations of delusion can help us to better assess epistemic goodness (Bortolotti, 2015) and investigations of depression can help us to better understand the nature of belief (Lane & Flanagan, 2015). Similarly, investigations of rationality in nonhuman animals can assist with the drawing of important distinctions among types of rationality (Bermudez, 2003), while investigations of artificial intelligence can assist with understanding how knowledge relevance is or should be determined (Ford and Pylyshyn, 1996).

Of course in this broader sense we can only strive to approximate rationality, and the effort devoted to this task is carried out in the dark, so

to speak, since there is no mature consensus theory of approximation. A time-honored alternative to approximation of the ideal is avoidance of error, a view associated with Pearce's "self-correcting thesis" (Mayo, 2005), a thesis no less problematic than the attempt to approximate rationality's ideal. Nevertheless, human success at having bootstrapped ourselves from common sense to science shows that taking concepts like "approximation of rationality" and "avoidance of irrationality" as rough-and-ready goals can drive progress.

But how does progress occur when the goal is inexact? Consider that every schoolboy learns to identify "north" in a rough-and-ready way by facing the sun before noontime and extending his left arm to the left side; a slightly more sophisticated technique, provided one is in the Northern Hemisphere, is to identify Polaris at Ursa Minor's handle tip. But neither of these will indicate true or geodesic north with precision; for that, surveying techniques are required. Complicating matters even more, geodesic north is distinct from magnetic north, the direction pointed to by a compass. And there are yet more distinct meanings of "north." Nevertheless, people have been finding their way north, approximately, for millennia, with methods that any child can learn.

Approximating rationality is like heading north. Human beings had probably been tilting toward rationality since the Pleistocene, long before Aristotle's theory of the syllogism, a seminal development that led to centuries of advances in rational thought, including the mathematical logic of Frege and Russell. During the last half century the pace of progress has accelerated, albeit in a way that was unanticipated. The very idea that we are rational has been challenged by findings from cognitive science, psychology, and related disciplines. Flashpoints have been many: base rate neglect, availability cascade, conjunction fallacy, belief perseverance, and so on. A modest conclusion that can be derived from these experimental findings about how humans reason is that we need a context-sensitive, carefully nuanced reassessment of rationality; a radical conclusion, as folk and scientific or scholarly models of rationality are deeply confused. This volume takes the former, not the latter, as its premise. Humans are predictably "irrational" at times, when dealing with some kinds of problems, when impaired, and when in certain contexts. But by attending closely to those varied contexts, and informed by a half century of scientific studies and philosophical efforts to achieve reflective equilibrium, the study of rationality can be rejuvenated, even applied to sober attempts at assessing irrationality.

Rationality cannot be approached asymptotically, not literally that is. Nevertheless, we can make more satisfactory our precisifications of "rationality" or its approximation. By aiming in its general direction, through analyses of some diverse contexts in which it is realized, while attending to various ways in which it is modulated, the project of understanding

rationality can forge ahead. In this volume we concentrate on diverse but not discrete contexts: scientific, communicative, pathological, nonhuman, ostensible irrationality in Chinese philosophy, and the modulation of reason by emotion. This scope and variety is well suited to this early postrationality wars era. An important step in reassessing "rationality" is to situate it in variegated contexts, so to better understand how belief-forming processes are relativized and constrained. Future models of rationality, whether descriptive or normative, will need to be grounded in just such an expansive foundation, if they are to further advance this seminal project.

Acknowledgments

For the generous giving of their time to detailed discussions of rationality and its constraints, I express heartfelt gratitude to Reinhard Selten as well as to all participants in the Conference on Reason and Rationality, hosted by Academia Sinica's Institute of European and American Studies. And, for discussions of these and related matters from days passed but not forgotten, the author is also grateful to Herbert Simon, Paul Lazarsfeld, Carl Hempel, and Tony O'Dea. Funding for this research was, in part, provided by National Science Council (Ministry of Science and Technology) of Taiwan Research Grants 100-2410-H-038-009-MY3, 102-2420-H-038-001-MY3, and 104-2420-H-038-001-MY3.

References

Bermudez, J. L. (2003). *Thinking without words*. New York: Oxford University Press.

Bortolotti, L. (2010). *Delusions and other irrational beliefs*. New York: Oxford University Press.

Bortolotti, L. (2015). The epistemic innocence of motivated delusions. *Consciousness and Cognition*, *33*, 490–499.

Botterill, G., & Carruthers, P. (1999). *The philosophy of psychology*. New York: Cambridge University Press.

Byron, M. (Ed.). (2004). *Satisficing and maximizing: Moral theorists on practical reason*. New York: Cambridge University Press.

Churchland, P. S. (2002). *Brain-wise: Studies in neurophilosophy*. Cambridge, MA: The MIT Press.

Clark, A. (2001). Reasons, robots and the extended mind. *Mind and Language*, *16*(2), 121–145.

Dahan-Katz, L. (2013). The implications of heuristics and biases research on moral and legal responsibility: A case against the reasonable person standard. In N. A. Vincent (Ed.), *Neuroscience and legal responsibility* (pp. 135–161). New York: Oxford University Press.

Damasio, A. R. (1999). *The feeling of what happens*. New York: Harcourt Brace.

Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford: Clarendon Press.

Ford, K. M., & Pylyshyn, Z. W. (Eds.). (1996). *The robot's dilemma revisited: The frame problem of artificial intelligence*. New York: Praeger.

Frith, C. D. (1998). Deficits and pathologies. In W. Bechtel, & G. Graham (Eds.), *A companion to cognitive science* (pp. 380–390). Malden, MA: Blackwell.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Glymour, C. (1992). *Thinking things through*. Cambridge, MA: The MIT Press.

Hamm, R. M. (1996). Physicians neglect base rates, and it matters. *Behavioral and Brain Sciences*, *19*, 25–26.

Hempel, C. G. (1965). *Aspects of scientific explanation: And other essays in the philosophy of science*. New York: The Free Press.
Hohwy, J. (2013). *The predictive mind*. Oxford, UK: Oxford University Press.
Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment and uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
Kahneman, D., & Tversky, A. (1972). Subjective probability. *Cognitive Psychology*, *3*, 430–454.
Kahneman, D., & Tversky, A. (1982). On the study of statistical intuition. *Cognition*, *11*, 123–142.
Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge: Cambridge University Press.
Karni, E. (2014). Axiomatic foundations of expected utility and subjective probability. In M. Machina, & K. Viscusi (Eds.), *Handbook of the economics of risk and uncertainty* (pp. 1–39). Amsterdam: Elsevier.
Kuhn, T. (2000a). Rationality and theory choice. In J. Conant, & J. Haugeland (Eds.), *The road since Structure* (pp. 208–215). Chicago, IL: The University of Chicago Press.
Kuhn, T. (2000b). Reflections on my critics. In J. Conant, & J. Haugeland (Eds.), *The road since Structure* (pp. 123–175). Chicago, IL: The University of Chicago Press.
Ladner, R. E. (1975). On the structure of polynomial time reducibility. *Journal of the Association of Computing Machinery*, *22*, 155–171.
Lane, T., & Flanagan, O. (2015). Neuroexistentialism, eudaimonics, and positive illusions. In B. Kaldis (Ed.), *Mind and society: Cognitive science meets the social sciences*. Synthese Library Series: Studies in Epistemology, Logic, Methodology, and Philosophy of Science. Dordrecht, Netherlands: Springer.
Lazarsfeld, P. F., & Oberschall, A. R. (1965). Max Weber and empirical social research. *American Sociological Research*, *31*(2), 185–199.
Mayo, D. G. (2005). Peircean induction and the error-correcting thesis. *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy*, *41*(2), 299–319.
Midgley, M. (2001). *Science and poetry*. London: Routledge & Kegan Paul.
Morse, S. J. (2007). New neuroscience, old problems: Legal implications of brain science. In W. Glannon (Ed.), *Defining right and wrong in brain science* (pp. 195–205). Washington, DC: Dana Press.
Newton-Smith, W. H. (1981). *The rationality of science*. London: Routledge & Kegan Paul.
Nisbett, R., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, *32*(5), 932–943.
Nisbett, R., & Ross, L. (1980). *Human inference*. Englewood Cliffs, NJ: Prentice-Hall.
Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.
Piattelli-Palmarinia, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: John Wiley & Sons.
Quine, W. V. (1976). *The ways of paradox and other essays, Revised and enlarged edition*. Cambridge, MA: Harvard University Press.
Rescher, N. (1988). *Rationality: A philosophical inquiry into the nature and rationale of reason*. New York: Oxford University Press.
Richardson, R. C. (1998). Heuristics and satisficing. In W. Bechtel, & G. Graham (Eds.), *A companion to cognitive science* (pp. 566–575). Malden, MA: Blackwell Publishers.
Samuels, R., Stich, S., & Bishop, M. (2002). Ending the rationality wars: How to make disputes about human rationality disappear. In R. Elio (Ed.), *Common sense, reasoning, and rationality* (pp. 236–268). New York: Oxford University Press.
Sellars, W. (1962). *Empiricism and the philosophy of mind*. London: Routledge & Kegan Paul.
Selten, R. (2001). What is bounded rationality? In G. Gigerenzer, & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 13–36). Cambridge, MA: The MIT Press.
Simon, H. (1972). Theories of bounded rationality. In C. McGuire, & R. Radner (Eds.), *Decision and organization* (pp. 161–176). Amsterdam: North-Holland.

Simon, H. (1983). *Reason in human affairs*. Stanford, CA: Stanford University Press.
Simon, H. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: The MIT Press.
Stanovich, K. E. (2005). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago, IL: The University of Chicago Press.
Stanovich, K. E. (2010). *Rationality and the reflective mind*. New York: Oxford University Press.
Trigg, R. (1993). *Rationality and science: Can science explain everything?* Cambridge, MA: Blackwell.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.
von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
Weber, M. (1978). *Economy and society: An outline of interpretive sociology*. Oakland, CA: University of California Press.