17 Individual differences as essential components of heuristics and biases research

Keith E. Stanovich
University of Toronto

Richard F. West
James Madison University

Maggie E. Toplak
York University, Toronto

Introduction

We are honoured to contribute to this volume because Jonathan Evans' work has been a key inspiration to us since even before we began contributing to the thinking and reasoning literature. Richard and Keith had been admirers of the heuristics and biases tradition from its inception in the early 1970s. However, their first research contributions were in the psychology of reading and this occupied them for 15 years (see Stanovich, 2000; Stanovich, Cunningham, & West, 1998). By the late 1980s though, we had decided to make a contribution to the literature on thinking and reasoning that we had admired so much for so long. Keith set off to Cambridge in early 1991 on a sabbatical with reading colleague Usha Goswami, and Rich visited for an important brainstorming week. The sabbatical really ended up being about reasoning and rationality rather than reading, however, and there were two books that we took to England to study in detail: Stich's *The Fragmentation of Reason* (1990) and Jonathan's *Biases in Human Reasoning* (1989). The former was obviously to bone up on the philosophical issues surrounding the concept of rational thought. However, the latter really became our bible on that sabbatical. In an amazingly few pages Jonathan zeroed in on the key issues surrounding the important biases and also gave the relative novice an introduction to the central tasks in the literature. That book, his insights as one of the first dual-process theorists, and almost four decades of continuous creative work in the area have made Jonathan one of our most important intellectual guideposts throughout our careers in this field.

The Great Rationality Debate

Our initial contribution to the field was to argue for the potential usefulness of individual difference data in the disputes surrounding what has been termed the Great Rationality Debate in cognitive science. It concerns the interpretation of the voluminous data indicating that human behaviour deviates from optimal standards as determined by decision scientists. How to interpret these deviations is, however, a matter of contentious dispute. As Tetlock and Mellers (2002) point out, "the debate over human rationality is a high-stakes controversy that mixes primordial political and psychological prejudices in combustible combinations" (p. 97). The so-called Great Debate about human rationality is a "high stakes controversy" because it involves nothing less than the models of human nature that underlie economics, moral philosophy, and the personal theories (folk theories) we use to understand the behaviour of other humans.

The root of the debate is the substantial research literature – one comprising literally hundreds of empirical studies conducted for more than three decades – showing that people's responses often deviate from the performance traditionally considered normative on many reasoning tasks. For example, people assess probabilities incorrectly; they display confirmation bias; they test hypotheses inefficiently; they violate the axioms of utility theory; they do not properly calibrate degrees of belief; they overproject their own opinions onto others; they allow prior belief to become implicated in their evaluation of evidence and arguments; and they display numerous other information processing biases (for summaries of the large literature, see Baron, 2008; Camerer, Loewenstein, & Rabin, 2004; Evans, 1989, 2007a; Gilovich, Griffin, & Kahneman, 2002; Johnson-Laird, 2006; Kahneman & Tversky, 2000; Koehler & Harvey, 2004; LeBoeuf & Shafir, 2005; Nickerson, 2008; Pohl, 2004; Stanovich, 1999, 2009). Indeed, demonstrating that descriptive accounts of human behaviour diverged from normative models was a main theme of the so-called heuristics and biases literature from its beginning in the 1970s and early 1980s (Dawes, 1976; Evans, 1984, 1989; Kahneman & Tversky, 1972, 1973; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1973, 1974; Wason & Evans, 1975).

The interpretation of the gap between descriptive models and normative models in the human reasoning and decision-making literature has been the subject of contentious debate for almost three decades now (Cohen, 1981, 1983; Evans & Over, 1996; Gigerenzer, 1996; Kahneman, 1981; Kahneman & Tversky, 1983, 1996; Koehler, 1996; Nickerson, 2008; Stanovich, 1999; Stein, 1996). The debate has arisen because some investigators wished to interpret the gap between the descriptive and the normative as indicating that human cognition was characterized by systematic irrationalities. Due to the emphasis that these theorists place on reforming human cognition, they were labelled the Meliorists by Stanovich (1999). Disputing this contention were numerous investigators (termed the Panglossians, see Stanovich, 1999) who argued that there were other reasons why reasoning might not accord with normative theory – reasons that prevent the ascription of irrationality to subjects. First, instances of reasoning might depart from normative standards due to performance errors – temporary lapses of attention, memory deactivation, and other sporadic information processing mishaps. Second, there may be stable and inherent computational limitations that prevent the normative response (Cherniak, 1986; Oaksford & Chater, 1993, 1995, 1998; Stich, 1990). Third, in interpreting performance, we might be applying the wrong normative model to the task (Koehler, 1996). Alternatively, we may be applying the correct normative model to the problem as set, but the subject might have construed the problem differently and be providing the normatively appropriate answer to a different problem (Adler, 1984, 1991; Berkeley & Humphreys, 1982; Broome, 1990; Hilton, 1995; Schwarz, 1996; Stanovich, 1999).

However, in referring to the various alternative explanations (other than systematic irrationality) for the normative/descriptive gap, Rips (1994) warns that "a determined skeptic can usually explain away any instance of what seems at first to be a logical mistake" (p. 393). In an earlier criticism of Henle's (1978) Panglossian position, Johnson-Laird (1983) made the same point: "There are no criteria independent of controversy by which to make a fair assessment of whether an error violates logic. It is not clear what would count as crucial evidence, since it is always possible to provide an alternative explanation for an error." (p. 26). The most humorous version of this argument was made by Kahneman (1981) in his dig at the Panglossians who seem to have only two categories of errors, "pardonable errors by subjects and unpardonable ones by psychologists" (p. 340). Referring to the four classes of alternative explanation discussed above – random performance errors, computational limitations, alternative problem construal, and incorrect norm application – Kahneman notes that Panglossians have "a handy kit of defenses that may be used if [subjects are] accused of errors: temporary insanity, a difficult childhood, entrapment, or judicial mistakes – one of them will surely work, and will restore the presumption of rationality" (p. 340).

These comments by Rips (1994), Johnson-Laird (1983), and Kahneman (1981) highlight the need for principled constraints on the alternative explanations of normative/descriptive discrepancies. We have tried to show that patterns of individual differences might help to provide just such principled constraints (Stanovich, 1999; Stanovich & West, 1998a, 1998b, 1998c, 1998d, 1999, 2000; West & Stanovich, 2003). For example, Panglossian theorists who argue that discrepancies between actual responses and those dictated by normative models are not indicative of human irrationality (e.g., Cohen, 1981) sometimes attribute the discrepancies to performance errors. Borrowing the idea of a competence/performance distinction from linguists (see Stein, 1996, pp. 8–9), these theorists view performance errors as the failure to apply a rule, strategy, or algorithm that is part of a person's competence because of a momentary and fairly random lapse in ancillary processes necessary to execute the strategy (lack of attention, temporary memory deactivation, distraction, etc.). Stein (1996) explains the idea of a performance error by referring to a "mere mistake" – a more colloquial notion that involves "a *momentary lapse*, a divergence from some typical behaviour. This is in contrast to attributing a divergence from norm to reasoning in accordance with principles that diverge from the normative principles of reasoning. Behaviour due to irrationality connotes a *systematic* divergence from the norm" (p. 8). Similarly, in the heuristics and biases literature, the term bias is reserved for systematic deviations from normative reasoning and does not refer to transitory processing errors ("a bias is a source of error which is systematic rather than random," Evans, 1984, p. 462). More technically, performance errors represent algorithmic-level problems that are transitory in nature. Nontransitory problems at the algorithmic level that would be expected to recur on a re-administration of the task should instead be viewed as computational limitations.

Another way to think of the performance error explanation is to conceive of it within the true score/measurement error framework of classical test theory. Mean or modal performance might be viewed as centered on the normative response – the response all people are trying to approximate. However, scores will vary around this central tendency due to random performance factors (error variance). In fact, a parallel argument has been made in economics where, as in reasoning, models of perfect market rationality are protected from refutation by positing the existence of local market mistakes of a transitory nature (temporary information deficiency, insufficient attention due to small stakes, distractions leading to missed arbitrage opportunities, etc.).

This notion of a performance error as a momentary attention, memory, or processing lapse that causes responses to appear non-normative even when competence is fully normative has implications for patterns of individual differences across reasoning tasks. For example, the strongest possible form of this view is that *all* discrepancies from normative responses are due to performance errors. This strong form of the hypothesis has the implication that there should be virtually no correlations among non-normative processing biases across tasks. If each departure from normative responding represents a momentary processing lapse due to distraction, carelessness, or temporary confusion, then there is no reason to expect covariance among biases across tasks (or covariance among items *within* tasks, for that matter) because error variances should be uncorrelated.

In contrast, positive manifold (uniformly positive bivariate associations in a correlation matrix) among disparate tasks in the heuristics and biases literature – and among items within tasks – would call into question the notion that all variability in responding can be attributable to performance errors. This was essentially Rips and Conrad's (1983) argument when they examined individual differences in deductive reasoning: "Subjects' absolute scores on the propositional tests correlated with their performance on certain other reasoning tests. . . . If the differences in propositional reasoning were merely due to interference from other performance factors, it would be difficult to explain why they correlate with these tests" (p. 282–283).

In several studies, we have found very little evidence for the strong version of the performance error view. With virtually all of the tasks from the heuristics and biases literature that we have examined, there is considerable internal consistency. Further, at least for certain classes of task, there are significant cross-task correlations (Stanovich & West, 1998c, 2000; West, Toplak, & Stanovich, 2008). Our purpose here is not to review the results of these studies, but to illustrate how individual differences shed light on a major dispute in the great rationality debate – how to interpret the normative/descriptive gap in empirical studies. As with performance errors, individual difference findings have implications for the argument that algorithmic limitations create discrepancies between descriptive and normative models. A strong correlation between measures of cognitive capacity and a rational thinking task suggests important algorithmic-level limitations that might make the normative response not prescriptive for those of lower cognitive capacity. In contrast, the absence of a correlation between the normative response and cognitive capacity suggests no computational limitation and thus no reason why the normative response should not be considered prescriptive. We have found that the computational limitations on most of the classic heuristics and biases tasks – at least as inferred from individual differences in cognitive ability – are not extreme.

More contentiously, however, we have argued that individual differences might be a piece of the puzzle in a very wide reflective equilibrium (Daniels, 1979, 1996; Elgin, 1996; Stich & Nisbett, 1980; Thagard, 1982; Thagard & Nisbett, 1983) regarding the wrong norm and alternative construal explanations of the normative/descriptive gap. Stanovich and West (2000) showed how Panglossian theorists use empirical data patterns to justify their critiques, but they rely exclusively on the modal response. An interesting aspect of some Panglossian positions is that because the descriptive is simply indexed to the normative, the latter can simply be "read off" from a competence model of the former (Cohen, 1982, terms this the norm extraction method). For example, Stein (1996) noted that this seems to follow from the Panglossian view because "whatever human reasoning competence turns out to be, the principles embodied in it are the normative principles of reasoning. . . . This argument sees the reasoning experiments as revealing human reasoning competence, and, thereby, as also revealing the norms. . . . The reasoning experiments just give us insight into what our reasoning abilities are; by knowing what these abilities are, we can determine what the normative principles of reasoning are" (p. 231–232).

Stein (1996) terms this type of Panglossian position the no extra-human norms view because it "rejects the standard picture of rationality and takes the reasoning experiments as giving insight not just into human reasoning competence but also into the normative principles of reasoning. . . . The no extra-human norms argument says that the norms just are what we have in our reasoning competence; if the (actual) norms do not match our preconceived notion of what the norms should be, so much the worse for our reconceived notions" (pp. 233–234). Stein (1996, p. 239) terms an extreme form of this strategy – that of explaining away *all* normative/descriptive gaps in terms of incorrect norm application – the "reject-the-norm strategy". It is noteworthy that this strategy is used exclusively by the Panglossian camp in the rationality debate, although this connection is not a necessary one.

Specifically, the reject-the-norm strategy is exclusively used to *eliminate* gaps between descriptive models of performance and normative models. When this type of critique is employed, the normative model that is suggested as a substitute for the rejected normative model is one that coincides perfectly with the descriptive model of the subjects' performance – thus preserving a view of human rationality as ideal. It is rarely noted that the strategy could be used in just the opposite way – to *create* gaps between the normative and descriptive. Situations where the modal response coincides with the standard normative model could be critiqued, and alternative models (normative models) could be suggested that would result in a new normative/descriptive gap. But this is never done. The Panglossian camp, often highly critical of empirical psychologists ("Kahneman and Tversky . . . and not their experimental subjects, commit the fallacies" Levi, 1983, p. 502), is never critical of psychologists who design reasoning tasks in instances where the modal subject gives the response the experimenters deem correct. Ironically, in these cases, according to the Panglossians, the same psychologists seem never to err in their task designs and interpretations.

It is quite clear that Cohen's (1979, 1981, 1986) trenchant criticisms of experimental psychologists would never have been written had human performance coincided with the standard normative models that the psychologists were using. The fact that the use of the reject-the-norm strategy is entirely contingent on the existence or nonexistence of a normative/descriptive gap suggests that the strategy is *empirically*, not conceptually, triggered (norms are never rejected for purely conceptual reasons when they coincide with the modal human response). What this means is that in an important sense the norms being endorsed by the Panglossian camp are conditioned (if not indexed entirely) by descriptive facts about human behaviour. Gigerenzer (1991) is clear about his adherence to an empirically-driven reject-the-norm strategy:

> Since its origins in the mid-seventeenth century. . . . When there was a striking discrepancy between the judgment of reasonable men and what probability theory dictated – as with the famous St. Petersburg paradox – then the mathematicians went back to the blackboard and changed the equations (Daston, 1980). Those good old days have gone. . . . If, in studies on social cognition, researchers find a discrepancy between human judgment and what probability theory seems

> to dictate, the blame is now put on the human mind, not the statistical model.
>
> (p. 109)

That Gigerenzer and Cohen concur here – even though they have somewhat different positions on normative justification – simply shows how widespread is the acceptance of the principle that descriptive facts about human behaviour condition our notions about the appropriateness of the normative models used to evaluate behaviour.

Interestingly, the descriptive component of performance around which Panglossian theorists almost always build their competence models (which index the normative in their view) is the central tendency of the responses (usually the mean or modal performance tendency). But if we are going to "read off" the normative from the descriptive in this way, why is this the only aspect of group performance that is relevant? Do the pattern of responses around the mode tell us anything? What about the rich covariance patterns that would be present in any multivariate experiment? Are these totally superfluous – all norm-relevant behavioural information residing in the mode? We have argued in several papers that if something about the normative must be inferred from the descriptive, then there is more information available than has traditionally been relied upon.

How should we interpret situations where the majority of individuals respond in ways that depart from the normative model applied to the problem by reasoning experts? Thagard (1982) calls the two different interpretations the populist strategy and the elitist strategy: "The populist strategy, favored by Cohen (1981), is to emphasize the reflective equilibrium of the average person. . . . The elitist strategy, favored by Stich and Nisbett (1980), is to emphasize the reflective equilibrium of experts" (p. 39). Thus, Thagard (1982) identifies the populist strategy with the Panglossian position and the elitist strategy with the Meliorist position.

But there are few controversial tasks in the heuristics and biases literature where all untutored laypersons disagree with the experts. There are always some who agree. Thus, the issue is not the untutored average person versus experts (as suggested by Thagard's formulation), but experts plus some laypersons versus other untutored individuals. Might the cognitive characteristics of those departing from expert opinion have implications for which normative model we deem appropriate? Larrick, Nisbett, and Morgan (1993) made just such an argument in their analysis of what justified the cost–benefit reasoning of microeconomics: "Intelligent people would be more likely to use cost-benefit reasoning. Because intelligence is generally regarded as being the set of psychological properties that makes for effectiveness across environments . . . intelligent people should be more likely to use the most effective reasoning strategies than should less intelligent people" (p. 333). Larrick et al. (1993) are alluding to the fact that we may want to condition our inferences about appropriate norms based not only on what response the majority of people make but also on what response the most cognitively competent subjects make.

Slovic and Tversky (1974) made essentially this argument years ago, although it was couched in very different terms in their paper and thus was hard to discern. Slovic and Tversky (1974) argued that descriptive facts about argument endorsement should condition the inductive inferences of experts regarding appropriate normative principles. In response to the argument that there is "no valid way to distinguish between outright rejection of the axiom and failure to understand it" (p. 372), Slovic and Tversky observed that "the deeper the understanding of the axiom, the greater the readiness to accept it" (pp. 372–373). Slovic and Tversky (1974) argued that this understanding/acceptance congruence suggested that the gap between the descriptive and normative was as a result of an initial failure to fully process and/or understand the task.

Slovic and Tversky's argument has been termed (see Stanovich & West, 1999) the understanding/acceptance assumption – that more reflective and engaged reasoners are more likely to affirm the appropriate normative model for a particular situation. From their understanding/acceptance principle, it follows that if greater understanding resulted in more acceptance of the axiom, then the initial gap between the normative and descriptive would be attributed to factors that prevented problem understanding (for example lack of ability or reflectiveness on the part of the subject). Such a finding would increase confidence in the normative appropriateness of the axioms and/or in their application to a particular problem. In contrast, if better understanding failed to result in greater acceptance of the axiom, then its normative status for that particular problem might be considered to be undermined. In short, the direction that performance moves in, in response to increased understanding, provides an empirical clue as to what is the proper normative model to be applied.

One might conceive of two generic strategies for applying the understanding/acceptance principle based on the fact that variation in understanding can be created or it can be studied by examining naturally occurring individual differences. Slovic and Tversky employed the former strategy by providing subjects with explicated arguments supporting the Allais or Savage normative interpretation, which we have also done with a variety of classic heuristic and biases tasks (Stanovich & West, 1999). Other methods of manipulating understanding have provided evidence in favour of some traditional normative principles. For example, it has been found that being forced to take more time or to provide a rationale for selections increases adherence to descriptive invariance (Miller & Fagley, 1991; Sieck & Yates, 1997; Takemura, 1992, 1993, 1994). Moshman and Geil (1998) found that group discussion facilitated performance on Wason's selection task.

As an alternative to the experimental manipulation of understanding, the understanding/acceptance principle can be transformed into an individual differences prediction. For example, the principle might be interpreted as indicating that more reflective, engaged, and intelligent reasoners are more likely to respond in accord with normative principles. Thus, it might be expected that those individuals with cognitive/personality characteristics more conducive to deeper understanding would be more accepting of the appropriate normative principles for a particular problem. This was the emphasis of Larrick et al. (1993) when they argued that more intelligent people should be more likely to use cost–benefit principles. Similarly, need for cognition – a dispositional variable reflecting the tendency toward thoughtful analysis and reflective thinking – has been associated with aspects of epistemic and practical rationality (Cacioppo, Petty, Feinstein, & Jarvis, 1996; Kardash & Scholes, 1996; Klaczynski, Gordon, & Fauth, 1997; LeBoeuf & Shafir, 2003; Smith & Levin, 1996; Verplanken, 1993).

The latter application of the understanding/acceptance principle derives from the assumption that a normative/descriptive gap that is disproportionately created by subjects with a superficial understanding of the problem provides no warrant for amending the application of standard normative models. It is the application (the individual differences application) that we have employed in most of our work. We have found that in most cases the traditionally used normative model was supported – but not always (Stanovich & West, 1998c, 2000, 2008b; West & Stanovich, 2003; West et al., 2008).

A framework for individual differences in heuristics and biases tasks

As we proceeded with our work in the 1990s – using individual differences to provide principled constraints on the interpretations of reasoning task responses – we began to focus on broader trends in our data. One was that the correlations between cognitive ability and performance on heuristics and biases tasks varied widely – from moderate to literally zero. What was obviously needed was a theoretical explanation that accounted for the variation in the magnitudes of the correlations.

One such attempt was made by Kahneman (2000) in his commentary on our summary of this work in *Behavioral and Brain Sciences*. He pointed out that many of the moderate correlations came from within-subjects designs that contain cues signalling the necessity of heuristic system override (Bartels, 2006; Fischhoff, Slovic, & Lichtenstein, 1979; Frisch, 1993; Kahneman & Tversky, 1982a; Shafir, 1998). Kahneman argued that between-subjects tests of the coherence of responses represents a much stricter criterion and perhaps a more appropriate one because "much of life resembles a between-subjects experiment" (Kahneman, 2000, p. 682). Shafir (1998) makes a similar argument when speculating about why people's behaviour is often at variance with their own normative intuitions. He argues that this discrepancy "mirrors a discrepancy between the nature of people's everyday experiences and the conditions that yield philosophical intuitions. In life, people typically experience and evaluate things one at a time, as in a between-subjects design, whereas many of the relevant intuitions result from concurrent, within-subject introspections" (p. 72).

That the mental factors operative in within-subjects designs might be different from those operative in between-subjects designs suggests that the individual difference factors associated with biased processing in the two different paradigms might also vary. LeBoeuf and Shafir (2003) have produced some data indicating that biases that are assessed within-subjects display different relationships with individual difference variables compared with biases assessed between-subjects. They found that various framing effects were associated with the need for cognition thinking disposition (see Cacioppo et al., 1996) when evaluated on a within-subjects basis but were independent of need for cognition when framing was assessed between subjects.

In a more systematic attempt to test Kahneman's conjecture, we found that correlations between cognitive ability and performance on heuristics and biases tasks tended to be low in between-subjects designs but not uniformly so (Stanovich & West, 2008b). On the left side of Table 17.1 is a list of heuristics and biases tasks that are dissociated from cognitive ability.[1] On the right side of Table 17.1 is a selection of heuristics and biases tasks that show significant correlations with cognitive ability. The between/within design structure of the tasks is a factor in explaining the nature of the relationship (associations tend to be lower in between-subjects tasks) but they are hardly the whole story. In a 2008 paper in the *Journal of Personality and Social Psychology* (Stanovich & West, 2008b), we attempted to develop a more comprehensive explanation.

We did begin, however, by building on Kahneman's insights. His argument begins with the distinction between coherence rationality and reasoning rationality. Reasoning rationality "requires an ability to reason correctly about the information currently at hand without demanding perfect consistency among beliefs that are not simultaneously evoked" (Kahneman & Frederick, 2005, p. 277). In contrast, "coherence is much stricter. . . . coherence requires choices and beliefs to be immune to variations of framing and context. This is a lot to ask for, but an inability to pass between-subjects tests of coherence is indeed a significant flaw" (Kahneman, 2000, p. 682). Kahneman and Frederick (2002; see Kahneman, 2000), utilizing a dual-process framework, argue that correlations with cognitive ability will occur only in the intermediate range of difficulty. There, they argue,

> intelligent people are more likely to possess the relevant logical rules and also to recognize the applicability of these rules in particular situations. In the terms of the present analysis, high-IQ respondents benefit from relatively efficient System 2 operations that enable them to overcome erroneous intuitions when adequate information is available. When a problem is too difficult for everyone, however, the correlation is likely to reverse.
>
> (Kahneman & Frederick, 2002, p. 68)

Table 17.1 Tasks in these experiments and in other studies that do and do not show associations with cognitive ability

Tasks/effects that fail to correlate with cognitive ability	*Tasks/effects that correlate with cognitive ability*
Noncausal base rate usage (Stanovich & West, 1998c, 1999, 2008b)	Causal base rate usage (Kokis et al., 2002; Stanovich & West, 1998c, 1998d)
Conjunction fallacy between-subjects (Stanovich & West, 2008b)	Outcome bias between- and within-subjects (Stanovich & West, 1998c, 2008b)
Framing between-subjects (Stanovich & West, 2008b)	Framing within-subjects (Bruine de Bruin, et al. 2007; Frederick, 2005; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999)
Anchoring effect (Stanovich & West, 2008b)	Denominator neglect (Kokis et al., 2002; Stanovich & West, 2008b)
Evaluability "Less is more" effect (Stanovich & West, 2008b)	Probability matching (West & Stanovich, 2003; Stanovich & West, 2008b)
Proportion dominance effect (Stanovich & West, 2008b)	Hindsight bias (Stanovich & West, 1998c)
Sunk cost effect (Stanovich & West, 2008b; Parker & Fischhoff, 2005)	Ignoring P(D/NH) (probability of the data given the negative hypothesis) (Stanovich & West, 1998d, 1999)
Risk/benefit confounding (Stanovich & West, 2008b)	Covariation Detection (Sá, West & Stanovich, 1999; Stanovich & West, 1998c, 1998d)
Omission bias (Stanovich & West, 2008b)	Belief bias in syllogistic reasoning (Macpherson & Stanovich, 2007; Stanovich & West, 1998c, 2008b)
One-side bias, within-subjects (Stanovich & West, 2008a)	Belief bias in modus ponens (Stanovich & West, 2008b)
Certainty effect (Stanovich & West, 2008b)	Informal argument evaluation (Stanovich & West, 1997, 2008b)
Willingness to pay/willingness to accept Difference (Stanovich & West, 2008b)	Four-card selection task (Stanovich & West, 1998a, 2008b; Toplak & Stanovich, 2002; Valentine, 1975)
Myside Bias – between- and within-subjects (Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a, 2008b; Toplak & Stanovich, 2003)	Expected value maximization in gambles (Benjamin & Shapiro, 2005; Frederick, 2005)
Newcomb's problem (Stanovich & West, 1999; Toplak & Stanovich, 2002)	Overconfidence effect (Bruine de Bruin et al., 2007; Stanovich & West, 1998c)

The phrase "possess the relevant logical rules and also to recognize the applicability of these rules in particular situations" suggests two conditions that have to be fulfilled for a heuristically based response to be overridden by analytic processing (Evans, 2003, 2006, 2007a, 2009; Kahneman & Frederick, 2002; Stanovich, 1999). These two conditions are actually the two sources of judgemental error that Kahneman and Tversky (1982a), two decades ago, labelled as: errors of application and errors of comprehension. The latter refers to errors that occur because people do not recognize the validity of a norm that they have violated because this knowledge has not been learned. The former occurs when the person fails to apply a rule they have learned.

In the remainder of this paper, we will use two slightly different terms for the loci of these problems. An error of comprehension we call a *mindware gap* (Stanovich, 2009; Stanovich, Toplak, & West, 2008). This is because in dual-process models, an important function of the analytic system is to take early representations triggered by the autonomous (heuristic) system offline and to substitute better responses. Mindware, a term coined by Perkins (1995; Clark, 2001, uses it in a slightly different way from Perkins' original coinage), refers to the rules, procedures, and strategies that can be retrieved by the analytic system and used to substitute for the heuristic response. However, if the mindware available to the analytic system for heuristic override has not been learned, then we have a case of a mindware gap.

In contrast, errors of application can only occur when the relevant mindware has been learned and is available for use in the override process. Errors of application occur when people fail to detect the situational cues indicating that the heuristically primed response needs to be overridden and an analytically derived response substituted. We give this requirement the label *override detection* (detecting the necessity for heuristic override). The above quote from Kahneman and Frederick (2002) suggests that cognitive ability differences only arise when the experimental task allows for variation in the presence of the relevant mindware and in the override detection process. It will be argued here that this analysis ignores a third potent source of non-normative responding that might be an even more important source of individual differences.

To understand our model of individual differences it is important to note that most of the tasks in the heuristics and biases literature were deliberately designed to pit a heuristically triggered response against a normative response. As Kahneman (2000) notes, "Tversky and I always thought of the heuristics and biases approach as a two-process theory" (p. 682). However, what this means is that even after the necessity for override has been detected and the relevant mindware is available, the conflict has to be resolved. Resolving the conflict in favour of the analytic response may require cognitive capacity, especially if cognitive decoupling (that is, inhibiting the heuristic response and simulating alternative responses) must take place for a considerable period of time while the analytic response is computed. Recent work on inhibition and executive functioning has indicated that such cognitive decoupling is very capacity demanding and that it is strongly related to individual differences in fluid intelligence (Duncan et al., 2008; Engle, 2002; Gray, Chabris, & Braver, 2003; Kane & Engle, 2002, 2003; Salthouse, Atkinson, & Berish, 2003; Unsworth & Engle, 2005, 2007).

It will be argued here that it is this third factor present in some heuristics and biases tasks – the necessity for sustained cognitive decoupling – that is the major source of the variability in the association between cognitive ability and task performance that is displayed in Table 17.1. Building on the conjectures of Kahneman (2000) and Kahneman and Frederick (2002), our framework for conceptualizing individual differences in heuristics and biases tasks is displayed in Figure 17.1. The question addressed in the first stage of the framework is whether, for a given task, the mindware is available to carry out override (whether the procedures and declarative knowledge are available to substitute an analytic response for a heuristic one). If the relevant mindware is not available, then the person must, of necessity, respond heuristically. It is immaterial whether the person detects the necessity for override or has the capacity to sustain override if the normatively appropriate

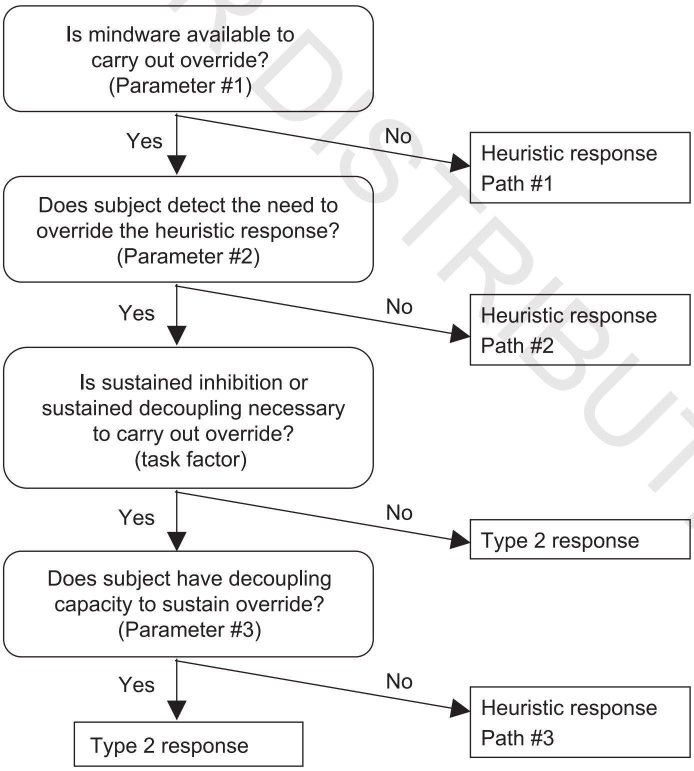

Figure 17.1 Framework for conceptualizing individual differences in heuristics and biases tasks.

response is simply not available. If the relevant mindware (probabilistic thinking skills, falsifiability tendencies, disposition to search for alternative explanations, sensitivity to contradiction, etc.) is not present, then participants will end up at what has been termed in Figure 17.1 Path #1 to a heuristic response.

If the relevant mindware is in fact available, then the next question that becomes operative is whether or not the person detects the need to override the heuristic response. Even if the relevant mindware is present, if the participant does not detect any reason to override the heuristic response, then it will be emitted (this is Path #2 to a heuristic response as labelled in Figure 17.1). Many heuristics and biases tasks lead people down this path. They do not detect the need to override the response that comes naturally (Kahneman, 2003) even though, in retrospect, they would endorse the norm that the heuristic response violated (Kahneman & Tversky, 1982a; Shafir, 1998; Shafir & Tversky, 1995; Thaler, 1987).

The next choice point in Figure 17.1 concerns the task rather than the participant. If the relevant mindware is present and if the need for override has been noted, the question then becomes whether or not the task requires sustained inhibition (cognitive decoupling) in order to carry out the override of the heuristic response. If not (or if the capacity required is low – this of course may not be an all or nothing issue), then the analytic response will be substituted for the heuristic response. In contrast, if the task requires sustained decoupling in order to carry out override, then we must ask whether the participant has the cognitive capacity that will be necessary. If so, then the analytic response will be given. If not, then the heuristic response will be given (Path #3 to the heuristic response in Figure 17.1) – despite the availability of the relevant mindware and the recognition of the need to use it.

In order for cognitive ability to associate with a bias, there must be differences correlated with cognitive ability at some of the choice points in the framework – that is, in some of the person parameters that branch toward or away from heuristic paths. As Kahneman (2000) notes "a task will be too difficult if (1) System 1 favors an incorrect answer, and (2) System 2 is incapable of applying the correct rule, either because the rule is unknown [mindware gap] or because the cues that would evoke it are absent [no override detection]" (p. 682). Performance on such a task will be floored and will show no association with cognitive ability. Some of the tasks in Table 17.1 are no doubt of this type (between-subjects conjunction effects for example). However, several of the tasks in Table 17.1 without associations with cognitive ability cannot be viewed as displaying floor effects. For a cognitive ability difference to be observed, there must be differential cleaving by intelligence at some of the critical nodes in Figure 17.1 – that is, there must be a correlation between intelligence and at least one of the person parameters. Of course, the partitioning of cognitive ability groups at each of the nodes will vary from task to task. We will advance here a generic conjecture about the source of associations with cognitive ability. The conjecture is that the primary source of associations with cognitive ability in heuristics and biases tasks is the way that people are partitioned by person parameter #3 ("does the person have the decoupling capacity to sustain override").

Cognitive decoupling in heuristics and biases tasks

There is direct evidence in the literature that intelligence tests (especially tests of fluid intelligence, see Carroll, 1993; Horn & Cattell, 1967; Horn & Noll, 1997) directly tap the ability to sustaining the decoupling of representations from the world so that cognitive simulations can be run which test the outcomes of imaginary actions (Currie & Ravenscroft, 2002; Dienes & Perner, 1999; Evans, 2007a; Evans & Over, 2004; Nichols & Stich, 2003). Thus, there is probably substantial variation in parameter #3 of the model because there is substantial variability in decoupling ability (Duncan, Emslie, Williams, Johnson, Freer, 1996; Kane & Engle, 2002; Salthouse et al., 2003). In contrast, we conjecture that, for many tasks in the heuristics and biases literature, the other two person parameters show only modest differential partitioning based on cognitive ability.

Regarding person parameter #1, it is true that the rules, knowledge, and strategies available to the analytic system to use in heuristic-system overrides are in part the product of past learning experiences. One might expect that people with more cognitive ability would profit more from learning experiences. However, the relevant mindware for our present discussion is not just generic procedural knowledge, nor is it the hodge-podge of declarative knowledge that is often used to assess crystallized intelligence on ability tests. Instead, it is a very special subset of knowledge related to: how one views probability and chance; whether one has the tools to think scientifically and the propensity to do so; the tendency to exhaustively examine possibilities; the tendency to avoid myside thinking; knowledge of some rules of formal and informal reasoning; and good argument evaluation skills. At least among the university students typically tested in these studies, acquiring these sets of skills and knowledge bases might be, experientially, very haphazard.

Although it is true that more intelligent individuals learn more things than the less intelligent, many thinking dispositions relevant to rationality are acquired rather late in life and the explicit teaching of this mindware is very spotty and inconsistent. For example, the tendency to think of alternative explanations for a phenomenon leads to the ability to more accurately infer causal models of events. Such principles are taught very inconsistently (by either explicit or implicit means). Or take, for example, the conjunction rule of probability, the violation of which is illustrated in the Linda Problem. Kahneman and Tversky (1982a) report that tests of rule endorsement and argument endorsement conducted after participants had made the conjunction error revealed that statistically sophisticated psychology graduate students did endorse the rule they had violated (they possessed the relevant mindware but did not detect the necessity for override). However, a majority of statistically naïve undergraduate students failed to endorse the conjunction rule – they lacked the relevant mindware ("much to our surprise, naive subjects did not have a solid grasp of the conjunction rule," Kahneman & Tversky, 1982a, p. 127). The lack of uniform teaching and learning conditions for the acquisition of this mindware might attenuate any natural correlation with intelligence that there would be if it were taught under uniform conditions.

Override detection (person parameter #2), we would argue, is perhaps even more likely to display a low correlation with cognitive ability. First, it would seem to be more of a thinking disposition (related to constructs like need for cognition, for instance, see Cacioppo et al., 1996) than a cognitive capacity. Psychometricians have long distinguished typical performance situations from optimal (sometimes termed maximal) performance situations (Ackerman, 1994, 1996; Ackerman & Heggestad, 1997; Ackerman & Kanfer, 2004; Cronbach, 1949; Matthews, Zeidner, & Roberts, 2002). Typical performance situations are unconstrained in that no overt instructions to maximize performance are given, and the task interpretation is determined to some extent by the participant. In contrast, optimal performance situations are those where the task interpretation is determined externally (not left to the participant). The participant is instructed to maximize performance, and is told how to do so. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of rational thinking dispositions (need for cognition, actively open-minded thinking, reflectivity/impulsivity) are often assessed under typical performance conditions.

Override detection, particularly in between-subjects designs, is exercised under typical rather than optimal conditions. It thus parses, in terms of the structure of cognitive abilities (Ackerman & Kanfer, 2004; Baron, 1985, 2008; Matthews, Zeidner, & Roberts, 2002; Sinatra & Pintrich, 2003; Sternberg, 1997) with thinking dispositions rather than cognitive capacity measures such as intelligence. For these theoretical reasons, we think that person parameter #2 in the framework is less the source of associations with cognitive ability than is person parameter #3.

There are two other ways that the influence of parameter #2, as a generator of individual differences, becomes attenuated – essentially by floor effects (as Kahneman, 2000, argues), but also by ceiling effects. Certain tasks in between-subjects designs (perhaps anchoring problems or the Linda Problem) give so few cues to the possibility of heuristic – analytic conflict that this parameter is probably floored for most subjects. Conversely, the instructions in other tasks (belief bias assessed with syllogisms for example), and some situations in real life ("the salesperson is trying to sell to you – don't forget") are so explicit in calling attention to heuristic – analytic conflict that this parameter is probably near ceiling.

The case of belief bias in syllogistic reasoning is probably a good illustration of our argument that it is person parameter #3 – the decoupling capacity parameter – that is the primary generator of associations with cognitive ability in rational thinking tasks[2] (see De Neys, 2006a, 2006b). The mindware available to reason logically on these simple categorical syllogisms (person parameter #1) is probably pretty uniformly present in the sample of university students studied here (and in most studies in the reasoning literature). The procedures needed to reason through the syllogisms used in these studies (for example, the invalid syllogism: all A are B, all C are B, therefore all C are A) are within the mindware of the vast majority of even the students who are classified as of low-cognitive ability in our university sample. Additionally, as just mentioned, the instructions on this task probably ceiling out person parameter #2 – override detection. Recall that the instructions to the task sensitize the participants to potential conflict (between argument validity and the truth of argument components). Thus, person parameters #1 and #2 probably leave little room for any individual difference variable to associate with performance.

In contrast, the task does require sustained cognitive decoupling (De Neys, 2006b). In the "rose" syllogism for example (All flowers have petals; Roses have petals; therefore, Roses are flowers – which is invalid), participants must suppress the tendency to endorse a valid response because of the "naturalness" (see Kahneman, 2003) of the conclusion – roses are flowers. This response must be held in abeyance while reasoning procedures work through the partially overlapping set logic indicating that the conclusion does not necessarily follow and that the syllogism is thus invalid. The reasoning process may take several seconds of perhaps somewhat aversive concentration (see Botvinick, Cohen, & Carter, 2004; Glenberg, 1997; Kahneman, 1973; Navon, 1989) – seconds during which the tendency to foreclose the conflict by acceding to the natural tendency to affirm "roses are flowers" (by responding "valid") must be avoided. Such response suppression while reasoning is closely related to the inhibitory and conflict resolution processes being studied by investigators examining the construct of executive functioning (Baddeley, Chincotta, & Adlam, 2001; Botvinick et al., 2004; Kane, 2003; Salthouse et al., 2003). Individual differences in such inhibitory processes have been found to be strongly associated with individual differences in fluid intelligence.

We conjecture that many of the other tasks that do show associations with cognitive ability (second column of Table 17.1) are tasks that involve some type of inhibition and/or sustained cognitive decoupling. For example, in within-subjects tests of outcome bias (Stanovich & West, 1998c) the appearance of the second item gives a pretty clear signal to the participant that there is an issue of consistency in their responses to the two different forms – that is, the within-subjects design probably puts person parameter #2 at ceiling, thus insuring that it is not the source of any associations with cognitive ability that are obtained. Detecting the need for consistency is not the issue. Instead, the difficulty comes from the necessity of inhibiting the tendency to downgrade the decision in the negative outcome condition, despite its having a better rationale than the positive outcome decision. Even in the between-subjects version of this task, one group of participants – those getting the negative outcome version – is alerted to the potential conflict between the seemingly good reasons to have the operation and the shockingly bad outcome. Participants must suppress the desire to sanction the decision, decouple their knowledge of the outcome, and simulate (see Evans & Over, 1996, 2004; Nichols & Stich, 2003) what they would have thought had they not known the outcome. Indeed, this condition creates a situation similar to those of various "curse of knowledge" paradigms (see Birch, 2005; Camerer, Loewenstein, & Weber, 1989; Gilovich, Medvec, & Sativsky, 1998; Hinds, 1999; Keysar & Barr, 2002).

The "curse of knowledge" logic of the negative item in the outcome bias task is similar to that in hindsight bias paradigms (e.g., Christiansen-Szalanski & Williams, 1991; Fischhoff, 1975; Pohl, 2004); which have also shown associations with cognitive ability (Stanovich & West, 1998c). In hindsight paradigms, the marking of the correct response sensitizes every respondent to the potential conflict involved – between what you know now, versus what you would have known without the correct response being indicated. Thus again, parameter #2 must be at ceiling. However, there is a need for sustained decoupling in the task, so whatever association between bias and cognitive ability exists on the task (a modest one, see Stanovich & West, 1998c) is likely generated by individual differences in parameter #3.

Within-subjects framing paradigms probably have a similar logic. The appearance of the second problem surely signals that an issue of consistency is at stake (putting parameter #2 at ceiling) and virtually all of the university students in these studies have acquired the value of consistency (parameter #1 is also at ceiling). The modest cognitive ability associations that are generated by this task probably derive from lower cognitive ability participants who cannot suppress the attractiveness of an alternative response despite the threat to consistent responding that it represents – in short, from variation in parameter #3. In contrast, between-subjects framing situations probably drive parameter #2 to a very low value (few people recognize that there is a conflict to be resolved between a potentially different response to an alternative framing), thus eliminating associations with individual differences (in the manner suggested by Kahneman, 2000).

The logic of the Linda Problem is similar. Transparent, within-subjects versions are easier because they signal the conflict involved and the necessity for override. Such versions create at least modest associations with cognitive ability. In the between-subjects version, however, individual differences are eliminated entirely because this design obscures the heuristic–analytic conflict and puts parameter #2 at floor.

As a final example, consider the difference between causal and noncausal base rates[3] illustrated in Table 17.1. Noncausal base-rate problems trigger conflict detection in so few participants that parameter #2 is floored and hence cognitive ability differences are eliminated. In contrast, in a classic causal base-rate problem such as the Volvo versus Saab problem (see endnote 3, p. 383), where aggregate information is pitted against indicant information, the aggregate information has a causal relationship to the criterion behaviour. Thus, the aggregate information in causal base-rate scenarios clearly signals that there are two pieces of information in conflict, parameter #2 is near ceiling, and individual differences are determined largely by parameter #3 (the sustained decoupling parameter) which is, we conjecture, linked to individual differences in cognitive ability. Thus, causal, but not noncausal, base-rate problems show cognitive ability differences.

It is important to note that this interpretation does not contradict the results of De Neys (De Neys & Glumicic, 2008; De Neys, Vartanian, & Goel, 2008) who demonstrated that various implicit measures of performance on noncausal base-rate problems (decision latencies, unannounced recall, brain activation in the anterior cingulate) indicated that conflict between base rate and indicant information was detected and that when indicant information was overridden that inhibition areas of the brain were activated (lateral prefrontal cortex). Unlike the classic 70/30 lawyer/engineer problem of Kahneman and Tversky (1973), very extreme base rates were used in the De Neys work, for example: "In a study 1000 people were tested. Among the participants there were 5 engineers and 995 lawyers". These extreme numbers serve to draw attention to the base rate and move parameter #2 to ceiling from its relatively low level in the traditional 70/30 version of the problem. The problem is turned from one where the pitfall is override detection to one where the central task is to inhibit the stereotype that is automatically triggered and replace it with reliance on the extreme base rate. Thus, individual differences in these extreme base-rate problems would be determined largely by parameter #3 (the sustained decoupling parameter), and it is thus to be expected that inhibition areas of the brain would be activated on trials where successful override is achieved. Likewise, because this version of the paradigm results in a moderate to high value of parameter #2, it is expected under our model that various implicit measures (including brain activation) would indicate that conflict between base rate and indicant information was detected.

Dual-process models and the reflective mind

We have used dual-process models to understand patterns of performance in heuristics and biases tasks. Our conjecture is that the reason why cognitive ability is dissociated from so many rational thinking tasks is that intelligence tests are very incomplete indices of Type 2 processing because they fail to assess a level of cognitive control that is critical to rational thought. In order to explicate what we mean by this statement, we will briefly sketch the dual-process conception that we work from – not a novel conception, but instead a synthesis of the literally dozens of such views now extant in the literature (see Evans, 2003, 2006, 2007a, 2008, 2009; Evans & Over, 1996; Kahneman & Frederick, 2002, 2005; Sloman, 1996; Stanovich, 1999, 2004).

Like many theorists, we distinguish Type 1 (heuristic) processing from Type 2 (analytic) processing. The defining feature of Type 1 processing is its autonomy. Other features often associated with Type 1 processing (speed, for example) are characteristic, but not defining of Type 1 processing. Type 1 processes are termed autonomous because: (1) their execution is mandatory when the triggering stimuli are encountered; (2) they do not put a heavy load on central processing capacity (i.e., they do not require conscious attention); (3) they are not dependent on input from high-level control systems; and (4) they can operate in parallel without interfering with themselves or with Type 2 processing.

Type 1 processing operations are multifarious. In a previous publication (Stanovich, 2004) they are referred to as TASS – the autonomous set of systems. 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. There has been extensive research on each of the different kinds of Type 1 processing (e.g., Atran, 1998; Buss, 2005; Evans, 2003; Fodor, 1983; Lieberman, 2000, 2003; Ohman & Mineka, 2001; Willingham, 1998, 1999). Type 1 processes conjoin the properties of automaticity, quasi-modularity, and heuristic processing as these constructs have been variously discussed in cognitive science (e.g., Bargh & Chartrand, 1999; Barrett & Kurzban, 2006; Carruthers, 2006; Coltheart, 1999; Evans, 1984, 2006, 2008, 2009; Samuels, 2005, 2009; Shiffrin & Schneider, 1977; Sperber, 1994).

Type 1 processing, because of its computational ease, is a common processing default. Type 1 processes are sometimes termed the adaptive unconscious (Wilson, 2002) in order to emphasize that Type 1 processes accomplish a host of useful things – face recognition, proprioception, language ambiguity resolution, depth perception, etc. – all of which are beyond our awareness. Heuristic processing is a term often used for Type 1 processing – to connote that it is fast, automatic, computationally inexpensive, and does not engage in extensive analysis of all the possibilities.

Type 2 processing is nonautonomous. It is relatively slow and computationally expensive. Many Type 1 processes can operate at once in parallel, but Type 2 processing is largely serial. Type 2 processing is often language based, but it is not necessarily so. One of the most critical functions of Type 2 processing is to override Type 1 processing. This is sometimes necessary because Type 1 processing is designed to get you into the right ballpark when solving a problem or making a decision, but it is not designed for the type of fine-grained analysis called for in situations of unusual importance (financial decisions, fairness judgements, employment decisions, legal judgments, etc.).

Heuristic processing depends on benign environments (see Stanovich, 2009, Ch. 6). In this context, a benign environment means one that contains useful cues that can be exploited by various heuristics (for example, affect-triggering cues, vivid and salient stimulus components, convenient anchors). Additionally, for an environment to be classified as benign, it also must contain no other individuals who will adjust their behaviour to exploit those relying only on heuristics. In contrast, a hostile environment for heuristics is one in which there are no cues that are usable by heuristic processes. Another way that an environment can turn hostile for a heuristic processor is if other agents discern the simple cues that are being used 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).

All of the different kinds of Type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are irrational in a particular context if not overridden. For example, often humans act as cognitive misers (see Stanovich, 2009) by engaging in attribute substitution (Kahneman & Frederick, 2002) – the substitution of an easy-to-evaluate characteristic for a harder one even if the easier one is less accurate. For example, the cognitive miser will substitute the less effortful attributes of vividness or affect for the more effortful retrieval of relevant facts (Kahneman, 2003; Slovic & Peters, 2006; Wang, 2009). But when we are evaluating important risks – such as the risk of certain activities and environments for our children – we do not want to substitute vividness for careful thought about the situation. In such situations, we want to employ Type 2 override processing to block the attribute substitution of the cognitive miser.

In order to override Type 1 processing, Type 2 processing must display at least two related capabilities. One is the capability of interrupting Type 1 processing and suppressing its response tendencies. Type 2 processing thus involves inhibitory mechanisms of the type that been the focus of recent work on executive functioning (Hasher, Lustig, & Zacks, 2007; Miyake, Friedman, Emerson & Witzki, 2000; Zelazo, 2004). But the ability to suppress Type 1 processing gets the job only half done. Suppressing one response is not helpful unless there is a better response available to substitute for it. Where do these better responses come from? One answer is that they come from processes of hypothetical reasoning and cognitive simulation[4] that are a unique aspect of Type 2 processing. When we reason hypothetically, we create temporary models of the world and test out actions (or alternative causes) in that simulated world.

In order to reason hypothetically we must, however, have one critical cognitive capability – we must be able to prevent our representations of the real world from becoming confused with representations of imaginary situations. For example, when considering an alternative goal state different from the one we currently have, we must be able to represent our current goal and the alternative goal and to keep straight which is which. Likewise, we need to be able to differentiate the representation of an action about to be taken from representations of potential *alternative* actions we are trying out in cognitive simulations. But the latter must not infect the former while the mental simulation is being carried out. Otherwise, we would confuse the action about to be taken with alternatives that we were just simulating.

As many theorists have argued, a process of cognitive decoupling (see Stanovich, 2009) is necessary to ensure this. Through the process of cognitive decoupling, the analytic system is responsible for the ability to create temporary models of the world and test the outcomes of imaginary actions (Currie & Ravenscroft, 2002; Evans, 2007a; Evans & Over, 1996, 2004; Nichols & Stich, 2003; Sterelny, 2001). By taking early representations triggered by Type 1 processing offline and substituting better responses that have survived the cognitive selection process of simulation, the Type 2 processing exemplifies activities often labelled as executive or inhibitory control. Decoupling for the purpose of offline simulation is a cognitively demanding operation. The raw ability to sustain such simulations while keeping the relevant representations decoupled is one key aspect of the brain's computational power that is being assessed by measures of fluid intelligence. This is becoming clear from converging work on executive function and working memory that both display correlations with intelligence that are quite high. The high degree of overlap in individual differences on working memory and other executive functioning tasks and individual differences in intelligence is probably due to the necessity for sustained decoupling operations on all the tasks involved (Duncan et al., 2008; Gray et al., 2003; Kane, 2003; Kane & Engle, 2002, 2003; Kane, Hambrick, & Conway, 2005; Salthouse et al., 2003).

Our studies of individual differences have led us to the important conclusion that Type 2 processing needs to be understood in terms of two levels of processing – the algorithmic level and the reflective level. We can see this if we consider the logic of autonomous system override. Type 1 processing will determine the response unless overridden by the algorithmic mechanisms of the analytic system. But override itself is initiated by higher level control. That is, the algorithmic level of the analytic system is conceptualized as subordinate to higher-level goal states and epistemic thinking dispositions, some of which have been studied empirically (e.g., Cacioppo et al., 1996; Stanovich & West, 1997, 2007). These goal states and epistemic dispositions exist at what might be termed the reflective level of processing – a level containing control states that regulate behaviour at a high level of generality. Such high-level goal states are common in the intelligent agents built by artificial intelligence researchers (Franklin, 1995; Pollock, 1995; Sloman, 1993; Sloman & Chrisley, 2003).

In Figure 17.2, we have presented the tripartite proposal in a simple form. In the spirit of Dennett's (1997) book *Kinds of Minds*, we have labelled the traditional source of Type 1 processing the autonomous mind. The distinction between the algorithmic and reflective mind derives from conceptual and empirical distinctions in the study of individual differences. At the algorithmic level the concern is information processing efficiency. The cognitive psychologist works largely at this level when they show that human performance can be explained by positing certain information processing mechanisms

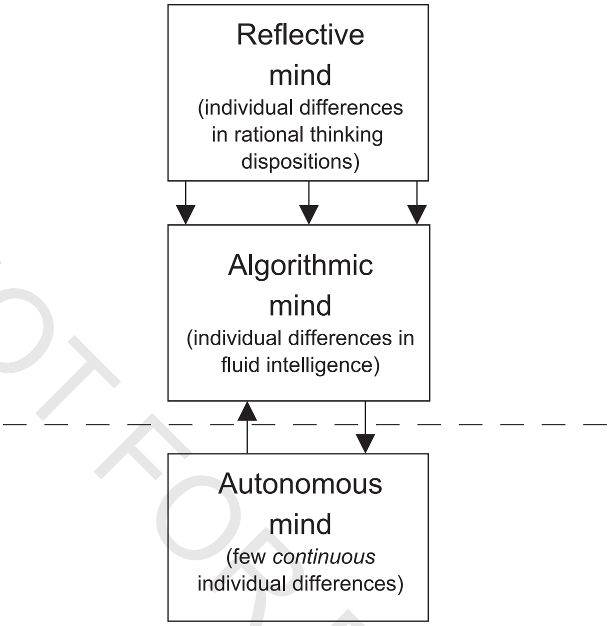

Figure 17.2 Individual differences in the tripartite structure.

in the brain (input coding mechanisms, perceptual registration mechanisms, short- and long-term memory storage systems, etc). We turn to the level of the reflective mind where we ask questions about the *goals* of the system's computations (*what* the system is attempting to compute and *why*). In short, the reflective mind is concerned with the goals of the system, beliefs relevant to those goals, and the choice of action that is optimal given the system's goals and beliefs. It is only at the level of the reflective mind where issues of rationality come into play. Importantly, the algorithmic mind can be evaluated in terms of efficiency but not rationality.

This concern for the efficiency of information processing as opposed to its rationality is mirrored in the status of intelligence tests. For example, investigators have attempted to decompose intelligence into more basic operations such as perceptual speed, discrimination accuracy, working memory capacity, and the efficiency of the retrieval of information stored in long-term memory (Ackerman, Kyllonen, & Richards, 1999; Carpenter, Just, & Shell, 1990; Deary, 2000, 2001; Hunt, 1987, 1999; Kane & Engle, 2002; Lohman, 2000; Sternberg, 1985, 1997, 2003; Unsworth & Engle, 2005). Such measures are indices of efficiency of information processing but not rationality – a point made clear by considering the distinction discussed earlier, that between typical performance situations and optimal (sometimes termed maximal) performance situations. Typical performance measures are measures of the reflective mind – they assess in part goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is determined externally. The person performing the task is instructed to maximize performance and is told how to do so. Thus, optimal performance measures examine questions of efficiency of goal pursuit – they capture the processing efficiency of the algorithmic mind. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical or rational thinking are often assessed under typical performance conditions.

The difference between the algorithmic mind and the reflective mind is captured in another well-established distinction in the measurement of individual differences – the distinction between cognitive ability and thinking dispositions. The former are, as just mentioned, measures of the efficiency of the algorithmic mind. The latter travel under a variety of names in psychology – thinking dispositions or cognitive styles being the two most popular. Many thinking dispositions concern beliefs, belief structure and, importantly, attitudes toward forming and changing beliefs. Other thinking dispositions that have been identified concern a person's goals and goal hierarchy. Examples of some thinking dispositions that have been investigated by psychologists are: actively open-minded thinking, need for cognition (the tendency to think a lot), consideration of future consequences, need for closure, superstitious thinking, and dogmatism (e.g., Ackerman & Heggestad, 1997; Baron, 1985, 2008; Cacioppo et al., 1996; Kruglanski & Webster, 1996; Perkins, 1995; Schommer, 1990; Sternberg, 2003; Sternberg & Grigorenko, 1997).

The literature on these types of thinking dispositions is vast and our purpose is not to review that literature here. It is only necessary to note that the types of cognitive propensities that these thinking disposition measures reflect are: the tendency to collect information before making up one's mind; the tendency to seek various points of view before coming to a conclusion; the disposition to think extensively about a problem before responding; the tendency to calibrate the degree of strength of one's opinion to the degree of evidence available; the tendency to think about future consequences before taking action; the tendency to explicitly weigh pluses and minuses of situations before making a decision; and the tendency to seek nuance and avoid absolutism. In short, individual differences in thinking dispositions are assessing variation in people's goal management, epistemic values, and epistemic self-regulation – differences in the operation of reflective mind. They are all psychological characteristics that underpin rational thought and action.

The cognitive abilities assessed on intelligence tests are not of this type. They are not about high-level personal goals and their regulation, or about the tendency to change beliefs in the face of contrary evidence, or about how knowledge acquisition is internally regulated when not externally directed. People have indeed come up with *definitions* of intelligence that encompass such things. Theorists often define intelligence in ways that encompass rational action and belief but, despite what these theorists argue, *the actual measures of intelligence in use assess only algorithmic-level cognitive capacity*.

No current intelligence test that is even moderately used in practice assesses rational thought or behaviour.

Individual differences in the tripartite structure

We can use this structure to understand why measures of intelligence provide a very incomplete assessment of Type 2 functioning. The algorithmic/reflective distinction reflects the contrast mentioned previously between optimal performance situations and typical performance situations. The psychology of individual differences maps this distinction into the difference between measures of cognitive capacity (intelligence) and measures of thinking dispositions.

Cognitive psychologists have focused on optimal performance indicators when studying the type of algorithmic-level cognitive capacities that underlie traditional psychometric intelligence (rightly so, because intelligence is assessed under maximal performance conditions). Fluid g (general intelligence) in the Horn–Cattell model (Carroll, 1993; Horn & Cattell, 1967) indexes primarily individual differences in cognitive capacity at the algorithmic level (as indicated in Figure 17.2). In contrast, thinking dispositions are measured with instructions emphasizing typical performance and are tapping the reflective mind. Many concern beliefs, belief structure and, importantly, attitudes toward forming and changing beliefs. Other cognitive styles concern a person's goals and goal hierarchy. Thinking disposition measures are telling us about the individual's goals and epistemic values – and they are indexing broad tendencies of pragmatic and epistemic self-regulation (Stanovich, 1999). They index individual differences in the reflective mind (as indicated in Figure 17.2). This is why such rational thinking dispositions will predict variance in performance on heuristics and biases tasks after the effects of general intelligence have been controlled (Bruine de Bruin, Parker & Fischhoff, 2007; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Newstead, Handley, Harley, Wright & Farrelly, 2004; Parker & Fischhoff, 2005; Sá & Stanovich, 2001; Stanovich & West, 1997, 1998c, 2000; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007; Toplak & Stanovich, 2002; West et al., 2008).

Figure 17.2 provides a framework for understanding why intelligence sometimes dissociates from thinking biases related to rationality. The reason is that rationality is a more encompassing construct than intelligence. Rationality is an organismic-level concept. It concerns the actions of an entity in its environment that serve its goals. As long as variation in thinking dispositions is not perfectly correlated with intelligence, then there is the statistical possibility of dissociations between rationality and intelligence. Substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Many different studies involving thousands of subjects have indicated that measures of intelligence display only moderate to weak correlations (usually less than

.30) with some thinking dispositions (e.g., actively open-minded thinking, need for cognition) and near zero correlations with others such as conscientiousness, curiosity, and diligence (Ackerman & Heggestad, 1997; Austin & Deary, 2002; Baron, 1982; Bates & Shieles, 2003; Cacioppo et al., 1996; Eysenck, 1994; Goff & Ackerman, 1992; Kanazawa, 2004; Kokis et al., 2002; Noftle & Robins, 2007; Reiss & Reiss, 2004; Zeidner & Matthews, 2000).

Integrative summary

In summary, what has been termed Type 2 processing in the dual-process literature (Evans, 2009) is composed of (at least) the distinct operations of the reflective and algorithmic minds, and only individual differences in the latter are strongly related to individual differences in intelligence. In terms of the framework illustrated in Figure 17.1, it is parameter #3 that relates strongly to the decoupling operation of the algorithmic mind and hence generates large individual differences based on cognitive ability. In contrast, it is conjectured that person parameter #1 is related more to the environmental/education history of the person that determines whether or not they have been exposed to the discrete packets of mindware that relate to rational thinking in the domains of cause, probability, choice tradeoffs, logic, and other specific domains. Finally, parameter #2 is posited to vary based upon individual differences in the reflective mind – specifically, whether the person has a tendency to default to heuristic responses or whether they are alert to cues indicating that heuristic override is necessary.

The framework illustrated in Figures 17.1 and 17.2 illustrates why rationality will not be uniformly related to intelligence. Instead, that relationship will depend upon the degree that rational responding requires sustained cognitive decoupling. When the heart of the task is recognizing the need for heuristic override but the override operation itself is easily accomplished, no sustained decoupling is necessary and rational thinking will depend more on the operations of the reflective mind than on the algorithmic mind. Thus, relationships with intelligence will be attenuated. Additionally, as Kahneman (2000) has argued, when detecting the necessity for override is very difficult (person parameter #2 is low), performance overall will be quite low and no relationships with cognitive ability will be evident.

Conversely however, highly intelligent people will display fewer reasoning biases when you tell them what the bias is and what they need to do to avoid it. That is, when parameters #1 and #2 are ceilinged and considerable cognitive capacity is needed to sustain decoupling while the correct response is computed, then highly intelligent people will do better in a rational thinking task. However, if there is no advance warning that biased processing must be avoided (as is the case in many between-subjects designs), then more intelligent individuals are not much more likely to perform any better on the task. Another way to phrase this is to say that, often, people of high-cognitive ability are no more likely to recognize the *need* for a normative principle than are individuals of lower-cognitive ability. When the former believe that nothing normative is at stake, they behave remarkably like other people (equally likely for example to be "anchored" into responding that redwoods are almost 1,000 feet tall! – as was the case in one of our experiments). If told, however, that they are in a situation of normative conflict and if resolving the conflict requires holding a prepotent response in abeyance, then the individual of high-cognitive ability will show less of many different cognitive biases.

Under the view presented here, there turn out to be numerous circumstances under which rational thought will dissociate from cognitive ability. First, cues to the necessity for override might be missing, thus leading most individuals to respond heuristically regardless of cognitive ability. Second, cues to the need for override might be present, but the disposition to respond to such cues may not be correlated with algorithmic capacity. If this is the case, there will be no association with cognitive ability as long as there is no need for sustained decoupling. The disposition to spot a heuristic–analytic conflict is not necessarily related to the computational power needed (in some situations) to resolve the conflict in favour of the analytic response. Thinking dispositions that relate to override cue detection are not assessed on intelligence tests and are thus not part of the intelligence construct.[5]

An important caveat to the model presented in Figure 17.1 is that which rational thinking tasks yield a conflict between heuristic and analytic responses is not fixed, but instead is a function of the individual's history of mindware acquisition. Early in developmental history, the relevant mindware will not be present and the heuristic response will be inevitable – no conflict will even be detected. Someone with no training in thinking probabilistically – or, for that matter, logically in terms of subset and superset – may experience no conflict in the Linda Problem. As experience with statistical and probabilistic thinking grows, a person will begin to experience more of a conflict because relevant mindware is available for use in the Type 2 simulation of an alternative response. The final developmental stage in this sequence might well be that the mindware used in analytic simulation becomes so tightly compiled that it is triggered from the autonomous mind in the manner of a natural heuristic response. 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. This final stage of processing is perhaps captured by developmental models of heuristic versus analytic processing that trace a trajectory where fluent adult performance looks very heuristic (Brainerd & Reyna, 2001; Ericsson & Charness, 1994; Klein, 1998; Reyna, Lloyd, & Brainerd, 2003; Reyna, Adam, Poirier, LeCroy & Brainerd, 2005).

Of course, with this discussion of what creates associations between biases and cognitive ability, we do not mean to draw attention away from one of our most salient outcomes – that a startlingly wide range of rational thinking tendencies appear to be somewhat independent of intelligence. Across many of the studies we have reported, we have found that many biases discussed in the heuristics and biases literature are surprisingly independent of cognitive ability. We say surprisingly because ever since Spearman (1904) first discovered positive manifold, intelligence indicators have correlated with a plethora of cognitive abilities and thinking skills that are almost too large to enumerate (e.g., Ackerman, Kyllonen, & Richards, 1999; Carroll, 1993; Deary, 2000, 2001; Deary, Whiteman, Starr, Whalley, & Fox, 2004; Lubinski, 2000, 2004; Lubinski & Humphreys, 1997). Nevertheless, as the left column of Table 17.1 indicates, the list of biases and effects that appear to be dissociated from intelligence is fairly long.

Of course, it is not true that all thinking biases are independent of cognitive ability. The right column of Table 17.1 lists some effects and biases from other studies where an association was found. Nevertheless, the correlations have tended to be modest: .35–.45 for belief bias in syllogistic reasoning, in the range of .25–.35 for various probabilistic reasoning tasks, in the range of .20–.25 for various covariation detection and hypothesis testing tasks, .25–.35 on informal reasoning tasks, .15–.20 with outcome bias measured within-subjects, .20–.40 with performance in the four-card selection task, .10–.20 with performance in various disjunctive reasoning tasks, .15–.25 with hindsight bias, .25–.30 with denominator neglect, and 05–.20 with various indices of Bayesian reasoning. All correlations were in the expected direction. Other investigators have found relationships of a similar effect size between cognitive ability and a variety of tasks in the heuristics and biases literature (Bruine de Bruin et al., 2007; DeShon, Smith, Chan, & Schmitt, 1998; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Klaczynski & Lavallee, 2005; Newstead et al., 2004; Parker & Fischhoff, 2005; Perkins & Ritchhart, 2004; Valentine, 1975).

The overall modest correlational trends are what led us to claim that intelligence tests are radically incomplete as measures of cognitive functioning (Stanovich, 2009; West et al., 2008). It is important to stress the word *cognitive* in this conclusion. Critics of intelligence tests are eager to point out that the tests ignore important parts of mental life – many largely noncognitive domains such as socioemotional abilities, empathy, and interpersonal skills, for example. However, a tacit assumption in such critiques is that although intelligence tests miss certain key noncognitive areas, they do encompass most of what is important in the cognitive domain. It is just this unstated assumption that we have challenged in our work – where we have shown that intelligence tests represent only a small sample of the cognitive skills importantly related to human rationality.

Notes

1 One caveat concerning the associations that we observed in these studies relates to the restriction of range in our sample. Certainly, it is true that individuals with average and above average cognitive ability are over represented in samples

composed entirely of university students. Nevertheless, the actual range in cognitive ability found among college students in the USA is quite large. In the past 30 years, the percentage of 25- to 29-years-olds in the USA who have attended college increased by 50%. By 2002, 58% of these young adults had completed at least one or more years of college, and 29% had received at least a bachelor's degree (US Department of Health and Human Services, 2003). However, the restriction of range in cognitive ability is somewhat greater in our sample, because our participants attended a moderately selective state university. The SAT Reasoning Test total means of our samples are roughly .60 of a standard deviation above the national mean of 1021 (College Board, 2006). The standard deviation of the distribution of scores in our sample is roughly .55–.70 of the standard deviation in the nationally representative sample.

2 Our conjecture here amounts to an endorsement of what Evans (2007b) calls the quality hypothesis regarding cognitive ability – that individuals higher in cognitive ability are more likely to compute the correct response *given* that they have engaged in Type 2 processing. The corresponding quantity hypothesis is that individuals higher in cognitive ability are more likely to engage in Type 2 processing.

3 Base rates that have a causal relationship to the criterion behaviour (Ajzen, 1977; Bar-Hillel, 1980, 1990; Tversky & Kahneman, 1979) are often distinguished from noncausal base rate problems – those involving base rates with no obvious causal relationship to the criterion behaviour. A famous noncausal problem is the well-known cab problem (see Bar-Hillel, 1980; Lyon & Slovic, 1976; Tversky & Kahneman, 1982):

> A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city in which the accident occurred. You are given the following facts: 85 percent of the cabs in the city are Green and 15 percent are Blue. A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each of the two colors 80 percent of the time. What is the probability that the cab involved in the accident was Blue?

(Amalgamating the base rate and the indicant according to Bayes' rule yields .41 as the posterior probability of the cab being blue.) The causal variant of the same problem substitutes for the first fact the phrase "Although the two companies are roughly equal in size, 85% of cab accidents in the city involve Green cabs and 15% involve Blue cabs" (Tversky & Kahneman, 1982; p. 157).

Another type of causal base rate problem is structured so that the participant has to make an inductive inference in a simulation of a real-life decision. The information relevant to the decision is conflicting and of two different types. One type of evidence is statistical: either probabilistic or aggregate base-rate information that favours one of the bipolar decisions. The other evidence is a concrete case or personal experience that points in the opposite direction. The classic Volvo versus Saab item (Fong, Krantz, & Nisbett, 1986, p. 285) provides an example. In this problem, a couple are deciding to buy one of two otherwise equal cars. Consumer surveys, statistics on repair records, and polls of experts favour the Volvo over the Saab. However, a friend reports experiencing a severe mechanical problem with the Volvo he owns. The participant is asked to provide advice to the couple. Preference for the Volvo indicates a tendency to rely on the large-sample information in spite of salient personal testimony. A preference for the Saab indicates reliance on the personal testimony over the opinion of experts and the large-sample information.

4 Hypothetical reasoning and cognitive simulation are central topics in cognitive science (see Barrett, Henzi, & Dunbar, 2003; Buckner & Carroll, 2007; Byrne, 2005;

Currie & Ravenscroft, 2002; Decety & Grezes, 2006; Dougherty, Gettys, & Thomas, 1997; Evans, 2007a; Evans & Over, 2004; Kahneman & Tversky, 1982b; Nichols & Stich, 2003; Oatley, 1999; Roese, 1997; Sterelny, 2001; Suddendorf & Corballis, 2007; Suddendorf & Whiten, 2001).

5 The Cognitive Reflection Test (CRT) developed by Frederick (2005) appears to be a combination index of the algorithmic mind and the reflective mind. It indexes both cognitive capacity and thinking dispositions, as exemplified by the tendency of the CRT to correlate with both intelligence indicators and thinking dispositions such as need for cognition and actively open-minded thinking, both in Frederick's (2005) own study and in unpublished work of our own.

Acknowledgements

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chairs program to Keith E. Stanovich; and also by a grant from the Social Sciences and Humanities Research Council of Canada to Maggie E. Toplak.

References

Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 4, pp. 1–27). Norwood, NJ: Ablex.

Ackerman, P. L. (1996). A theory of adult development: Process, personality, interests, and knowledge. *Intelligence*, *22*, 227–257.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Ackerman, P. L., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 119–141). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ackerman, P., Kyllonen, P., & Richards, R. (Eds.). (1999). *Learning and individual differences: Process, trait, and content determinants*. Washington, DC: American Psychological Association.

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, The Netherlands: Editions Rodopi.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, *35*, 303–314.

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *21*, 547–609.

Austin, E. J., & Deary, I. J. (2002). Personality dispositions. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 187–211). New Haven, CT: Yale University Press.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*, 641–657.
Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.
Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, *44*, 211–233.
Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Eds.), *Insights into decision making: A tribute to Hillel J. Einhorn* (pp. 200–216). Chicago, IL: University of Chicago Press.
Baron, J. (1982). Personality and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 308–351). Cambridge: Cambridge University Press.
Baron, J. (1985). *Rationality and intelligence*. Cambridge: Cambridge University Press.
Baron, J. (2008). *Thinking and deciding (Fourth Edition)*. Cambridge: Cambridge University Press.
Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*, 628–647.
Barrett, L., Henzi, P., & Dunbar, R. (2003). Primate cognition: From "what now?" to "what if?". *Trends in Cognitive Sciences*, *7*, 494–497.
Bartels, D. M. (2006). Proportion dominance: The generality and variability of favoring relative savings over absolute savings. *Organizational Behavior and Human Decision Processes*, *100*, 76–95.
Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: the relationship of inspection time and openness to g and Gc. *Intelligence*, *31*, 275–287.
Benjamin, D., & Shapiro, J. (2005, February 25). *Does cognitive ability reduce psychological bias*? JEL manuscript: J24, D14, C91.
Berkeley, D., & Humphreys, P. (1982). Structuring decision problems and the "bias heuristic". *Acta Psychologica*, *50*, 201–252.
Birch, S. A. J. (2005). When knowledge is a curse: Children's and adult's reasoning about mental states. *Current Directions in Psychological Science*, *14*, 25–29.
Botvinick, M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*, 539–546.
Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, pp. 41–100). San Diego, CA: Academic Press.
Broome, J. (1990). Should a rational agent maximize expected utility? In K. S. Cook & M. Levi (Eds.), *The limits of rationality* (pp. 132–145). Chicago, IL: University of Chicago Press.
Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*, 938–956.
Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49–57.
Buss, D. M. (Ed.). (2005). *The handbook of evolutionary psychology*. Hoboken, NJ: John Wiley.
Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, *97*, 1232–1254.

Camerer, C., Loewenstein, G., & Rabin, M. (Eds.). (2004). *Advances in behavioral economics*. Princeton, NJ: Princeton University Press.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404–431.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Carruthers, P. (2006). *The architecture of the mind*. New York: Oxford University Press.

Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.

Christiansen-Szalanski, J. J., & Williams, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*, 147–168.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press.

Cohen, L. J. (1979). On the psychology of prediction: Whose is the fallacy? *Cognition*, *7*, 385–407.

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

Cohen, L. J. (1982). Are people programmed to commit fallacies? Further thoughts about the interpretation of experimental data on probability judgment. *Journal for the Theory of Social Behavior*, *12*, 251–274.

Cohen, L. J. (1983). The controversy about irrationality. *Behavioral and Brain Sciences*, *6*, 510–517.

Cohen, L. J. (1986). *The dialogue of reason*. Oxford: Oxford University Press.

College Board (2006). College Board SAT: 2006 College-Bound Seniors: Total Group Profile Report. Retrieved February 26, 2010, from http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2006/national-report.pdf

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, *3*, 115–120.

Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.

Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. Oxford: Clarendon Press.

Daniels, N. (1979). Wide reflective equilibrium and theory acceptance ethics. *The Journal of Philosophy*, *76*, 256–282.

Daniels, N. (1996). *Justice and justification: Reflective equilibrium in theory and practice*. Cambridge: Cambridge University Press.

Daston, L. (1980). Probabilistic expectation and rationality in classical probability theory. *Historia Mathematica*, *7*, 234–260.

Dawes, R. M. (1976). Shallow psychology. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 3–11). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*, 130–147.

Decety, J., & Grezes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14.

De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly Journal of Experimental Psychology*, *59*, 1070–1100.

De Neys, W. (2006b). Dual processing in reasoning – Two systems but one reasoner. *Psychological Science*, *17*, 428–433.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299.

De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*, 483–489.

Dennett, D. C. (1997). Kinds of minds: Toward an understanding of consciousness. New York: Basic Books.

DeShon, R. P., Smith, M. R., Chan, D., & Schmitt, N. (1998). Can racial differences in cognitive test performance be reduced by presenting problems in a social context? *Journal of Applied Psychology*, *83*, 438–451.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.

Dougherty, M. R. P., Gettys, C. F., & Thomas, R. P. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes*, *70*, 135–148.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*, 257–303.

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., & Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology-General*, *137*, 131–148.

Elgin, C. Z. (1996). *Considered judgment*. Princeton, NJ: Princeton University Press.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.

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: Lawrence Erlbaum Associates Ltd.

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

Evans, J. St B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*, *13*, 378–395.

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

Evans, J. St B. T. (2007b). On the resolution of conflict in dual process theories of reasoning. *Thinking and Reasoning*, *13*, 321–339.

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. 33–54). Oxford: Oxford University Press.

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

Evans, J. St B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Eysenck, H. J. (1994). Personality and intelligence: Psychometric and experimental approaches. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 3–31). Cambridge: Cambridge University Press.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1979). Subjective sensitivity analysis. *Organizational Behavior and Human Performance*, *23*, 339–359.

Fodor, J. (1983). *Modularity of mind*. Cambridge: MIT Press.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253–292.

Franklin, S. (1995). *Artificial minds*. Cambridge, MA: MIT Press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.

Frisch, D. (1993). Reasons for framing effects. *Organizational Behavior and Human Decision Processes*, *54*, 399–429.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of Social Psychology*, *2*, 83–115.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, *103*, 592–596.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Gilovich, T., Medvec, V. H., & Sativsky, K. (1998). The illusion of transparency: Biased assessment of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, *75*, 332–346.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–55.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*, 537–552.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. St B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking and Reasoning*, *10*, 175–195.

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York: Oxford University Press.

Henle, M. (1978). Foreword. In R. Revlin, & R. E. Mayer (Eds.), *Human reasoning*. Washington DC: Winston.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, *118*, 248–271.

Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied*, *5*, 205–221.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 1–23.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. Flanagan, J. Genshaft, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York: Guilford Press.

Hunt, E. (1987). The next word on verbal ability. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 347–392). Norwood, NJ: Ablex.

Hunt, E. (1999). Intelligence and human resources: Past, present, and future. In P. Ackerman & P. Kyllonen (Eds.), *The future of learning and individual differences research: Processes, traits, and content* Washington, DC: American Psychological Association.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kahneman, D. (1981). Who shall be the arbiter of our intuitions? *Behavioral and Brain Sciences*, *4*, 339–340.

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. A. (2003). Perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

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., & Frederick, S. (2005). A model of intuitive judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267–293). New York: Cambridge University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1982a). On the study of statistical intuitions. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 493–508). Cambridge, MA: Cambridge University Press.

Kahneman, D., & Tversky, A. (1982b). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1983). Can irrationality be intelligently discussed? *Behavioral and Brain Sciences*, *6*, 509–510.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591.
Kahneman, D., & Tversky, A. (Eds.) (2000). *Choices, values and frames*. New York: Cambridge University Press.
Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review*, *111*, 512–523.
Kane, M. J. (2003). The intelligent brain in conflict. *Trends in Cognitive Sciences*, *7*, 375–377.
Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*, 637–671.
Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70.
Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.
Kardash, C. M., & Scholes, R. J. (1996). Effects of pre-existing beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology*, *88*, 260–271.
Keysar, B., & Barr, D. J. (2002). Self-anchoring in conversation: Why language users do not do what they "should". In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 150–166). New York: Cambridge University Press.
Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology*, *89*, 470–485.
Klaczynski, P. A., & Lavallee, K. L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology*, *92*, 1–24.
Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: adult age differences in everyday reasoning tasks. *Psychology and Aging*, *15*, 400–416.
Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
Koehler, D. J., & Harvey, N. (Eds.). (2004). Blackwell handbook of judgment and decision making. Oxford: Blackwell.
Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative and methodological challenges. *Behavioral and Brain Sciences*, *19*, 1–53.
Kokis, J, Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*, 26–52.
Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing the mind: "Seizing" and "freezing". *Psychological Review*, *103*, 263–283.
Larrick, R. P., Nisbett, R. E., & Morgan, J. N. (1993). Who uses the cost-benefit rules of choice? Implications for the normative status of microeconomic theory. *Organizational Behavior and Human Decision Processes*, *56*, 331–347.
LeBoeuf, R. A., & Shafir, E. (2003). Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making*, *16*, 77–92.

LeBoeuf, R. A., & Shafir, E. (2005). Decision making. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 243–266). Cambridge: Cambridge University Press.

Levi, I. (1983). Who commits the base rate fallacy? *Behavioral and Brain Sciences*, *6*, 502–506.

Lieberman, M. D. (2000). Intuiton: A social cognitive neuroscience approach. *Psychological Bulletin*, *126*, 109–137.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 44–67). New York: Cambridge University Press.

Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 285–340). Cambridge, MA: Cambridge University Press.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points". *Annual Review of Psychology*, *51*, 405–444.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "General Intelligence, Objectively Determined and Measured". *Journal of Personality and Social Psychology*, *86*, 96–111.

Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, *24*, 159–201.

Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica*, *40*, 287–298.

Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences*, *17*, 115–127.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science & myth*. Cambridge, MA: MIT Press.

Miller, P. M., & Fagley, N. S. (1991). The effects of framing, problem variations, and providing rationale on choice. *Personality and Social Psychology Bulletin*, *17*, 517–522.

Miyake, A., Friedman, N., Emerson, M. J., & Witzki, A. H. (2000). The utility and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking and Reasoning*, *4*, 231–248.

Navon, D. (1989). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system. *European Journal of Cognitive Psychology*, *1*, 191–238.

Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology*, *57A*, 33–60.

Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretense, self-awareness and understanding other minds*. Oxford: Oxford University Press.

Nickerson, R. S. (2008). *Aspects of rationality*. New York: Psychology Press.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*, 116–130.

Oaksford, M., & Chater, N. (1993). Reasoning theories and bounded rationality. In K. Manktelow & D. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 31–60). London: Routledge.

Oaksford, M., & Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning*, *1*, 121–152.

Oaksford, M., & Chater, N. (Eds.). (1998). *Rationality in an uncertain world*. Hove, UK: Psychology Press.

Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology*, *3*, 101–117.

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual differences approach. *Journal of Behavioral Decision Making*, *18*, 1–27.

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.

Perkins, D., & Ritchhart, R. (2004). When is good thinking? In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 351–384). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Pohl, R. (2004). Hindsight bias. In R. Pohl (Ed.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment and memory* (pp. 363–378). Hove, UK: Psychology Press.

Pollock, J. L. (1995). *Cognitive carpentry: A blueprint for how to build a person*. Cambridge, MA: MIT Press.

Reiss, S., & Reiss, M. M. (2004). Curiosity and mental retardation: Beyond IQ. *Mental Retardation*, *42*, 77–81.

Reyna, V. F., Adam, M. B., Poirier, K., LeCroy, C., & Brainerd, C. J. (2005). Risky decision making in childhood and adoelescence: A fuzzy-trace theory approach. In J. E. Jacobs & P. A. Klaczynski (Eds.), *The development of judgment and decision making in children and adolescents* (pp. 77–106). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Reyna, V. F., Lloyd, F. J., & Brainerd, C. J. (2003). Memory, development, and rationality: An integrative theory of judgment and decision making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 201–245). New York: Cambridge University Press.

Rips, L. J. (1994). *The logic of proof*. Cambridge, MA: MIT Press.

Rips, L. J., & Conrad, F. G. (1983). Individual differences in deduction. *Cognition and Brain Theory*, *6*, 259–285.

Roese, N. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*, 131–148.

Sá, W., Kelley, C., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: Individual differences in the coordination of theory and evidence. *Personality and Individual Differences*, *38*, 1149–1161.

Sá, W., & Stanovich, K. E. (2001). The domain specificity and generality of mental contamination: Accuracy and projection in judgments of mental content. *British Journal of Psychology*, *92*, 281–302.

Sá, W., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, *91*, 497–510.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594.

Samuels, R. (2005). The complexity of cognition: Tractability arguments for massive modularity. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (pp. 107–121). Oxford: Oxford University Press.

Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds. In J. St B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 129–146). Oxford: Oxford University Press.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*, 498–504.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahweh, NJ: Lawrence Erlbaum Associates, Inc.

Shafir, E. (1998). Philosophical intuitions and cognitive mechanisms. In M. R. DePaul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 59–83). Lanham, MD: Rowman & Littlefield Publishers, Inc.

Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), *Thinking* (Vol. 3, pp. 77–100). Cambridge, MA: The MIT 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.

Sieck, W., & Yates, J. F. (1997). Exposition effects on decision making: Choice and confidence in choice. *Organizational Behavior and Human Decision Processes*, *70*, 207–219.

Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sloman, A. (1993). The mind as a control system. In C. Hookway & D. Peterson (Eds.), *Philosophy and cognitive science* (pp. 69–110). Cambridge: Cambridge University Press.

Sloman, A., & Chrisley, R. (2003). Virtual machines and consciousness. *Journal of Consciousness Studies*, *10*, 133–172.

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

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*, 322–325.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science*, *19*, 368–373.

Smith, S. M., & Levin, I. P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, *9*, 283–290.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 39–67). Cambridge: Cambridge University Press.

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

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
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., Cunningham, A. E., & West, R. F. (1998). Literacy experiences and the shaping of cognition. In S. Paris & H. Wellman (Eds.), *Global prospects for education: Development, culture, and schooling* (pp. 253–288). Washington, DC: American Psychological Association.
Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. *Advances in child development and behavior*, *36*, 251–285.
Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357.
Stanovich, K. E., & West, R. F. (1998a). Cognitive ability and variation in selection task performance. *Thinking and Reasoning*, *4*, 193–230.
Stanovich, K. E., & West, R. F. (1998b). Individual differences in framing and conjunction effects. *Thinking and Reasoning*, *4*, 289–317.
Stanovich, K. E., & West, R. F. (1998c). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161–188.
Stanovich, K. E., & West, R. F. (1998d). Who uses base rates and P(D/~H)? An analysis of individual differences. *Memory & Cognition*, *26*, 161–179.
Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive Psychology*, *38*, 349–385.
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. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, *13*, 225–247.
Stanovich, K. E., & West, R. F. (2008a). On the failure of intelligence to predict myside bias and one-sided bias. *Thinking & Reasoning*, *14*, 129–167.
Stanovich, K. E., & West, R. F. (2008b). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*, 672–695.
Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.
Sterelny, K. (2001). *The evolution of agency and other essays*. Cambridge: Cambridge University Press.
Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.
Sternberg, R. J. (1997). *Thinking styles*. Cambridge: Cambridge University Press.
Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge: Cambridge University Press.
Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist*, *52*, 700–712.
Stich, S. P. (1990). *The fragmentation of reason*. Cambridge: MIT Press.
Stich, S. P., & Nisbett, R. E. (1980). Justification and the psychology of human reasoning. *Philosophy of Science*, *47*, 188–202.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel and is it unique to humans? *Behavioral and Brain Sciences*, *30*, 299–351.

Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals. *Psychological Bulletin*, *127*, 629–650.

Takemura, K. (1992). Effect of decision time on framing of decision: A case of risky choice behavior. *Psychologia*, *35*, 180–185.

Takemura, K. (1993). The effect of decision frame and decision justification on risky choice. *Japanese Psychological Research*, *35*, 36–40.

Takemura, K. (1994). Influence of elaboration on the framing of decision. *Journal of Psychology*, *128*, 33–39.

Tetlock, P. E., & Mellers, B. A. (2002). The great rationality debate. *Psychological Science*, *13*, 94–99.

Thagard, P. (1982). From the descriptive to the normative in philosophy and logic. *Philosophy of Science*, *49*, 24–42.

Thagard, P., & Nisbett, R. E. (1983). Rationality and charity. *Philosophy of Science*, *50*, 250–267.

Thaler, R. H. (1987). The psychology of choice and the assumptions of economics. In A. E. Roth (Ed.), *Laboratory experimentation in economics: Six points of view* (pp. 99–130). Cambridge: University of Cambridge Press.

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K. E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making*, *20*, 103–124.

Toplak, M., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, *94*, 197–209.

Toplak, M. E. & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology*, *17*, 851–860.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Tversky, A., & Kahneman, D. (1979). Causal schemas in judgements under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). Cambridge: Cambridge University Press.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*, *33*, 67–81.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.

US Department of Health and Human Services (2003). *Trends in the Well Being of America's Children & Youth, 2003*. US Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation, Retrieved February 26, 2010, from http://aspe.hhs.gov/hsp/ø3trends/

Valentine, E. R. (1975). Performance on two reasoning tasks in relation to intelligence, divergence and interference proneness: Content and context effects in reasoning. *British Journal of Educational Psychology*, *45*, 198–205.

Verplanken, B. (1993). Need for cognition and external information search: Responses to time pressure during decision-making. *Journal of Research in Personality*, *27*, 238–252.

Wang, L. (2009). Money and fame: Vividness effects in the National Basketball Association. *Journal of Behavioral Decision Making*, *22*, 20–44.

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

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition*, *31*, 243–251.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*, 930–941.

Willingham, D. T. (1998). A neuropsychological theory of motor-skill learning. *Psychological Review*, *105*, 558–584.

Willingham, D. T. (1999). The neural basis of motor-skill learning. *Current Directions in Psychological Science*, *8*, 178–182.

Wilson, T. D. (2002). *Strangers to ourselves*. Cambridge, MA: Harvard University Press.

Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 581–610). New York: Cambridge University Press.

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, *8*, 12–17.