Rational thinking involves adopting appropriate goals, taking the appropriate action given one’s goals and beliefs, and holding beliefs that are commensurate with available evidence. Traditional tests of intelligence are not good proxies for rational thinking skills because rational thought and intelligence are conceptually and empirically separable. Thus, the developmental trajectories of the former must be studied in their own right. Research has shown unequivocally that children do show the biases that have been displayed in the adult literature, but the developmental trends have been quite varied. Across the age ranges studied, there appear to be developmental increases in the avoidance of belief bias, analytic responding in the selection task, probabilistic reasoning, and reliance on statistical information in the face of conflicting personal testimonials. All of these trends are in the direction of increasingly rational thought, at least according to the most commonly employed normative standards. These findings contrast, however, with findings from developmental studies of myside bias and framing effects. Neither of these biases was attenuated by development. The lack of developmental decreases in these two biases is interestingly convergent with findings that neither bias displays much of a correlation with intelligence.

Reasoning skills have developmental trajectories that have been the subject of a moderate amount of research (Bjorklund, 2004; Goswami, 2004;
Because assessments of intelligence (and similar tests of cognitive ability) are taken to be the sine qua non of good thinking, it might be thought that such measures would serve as proxies for the developmental trajectories for judgment and decision-making skills. We argue in this chapter that such an assumption is misplaced. Traditional tests of intelligence are not good proxies for judgment and decision-making skills, and thus the developmental trajectories of the latter must be studied in their own right. Judgment and decision making are more properly regarded as components of rational thought, and it is often not recognized that rationality and intelligence (as traditionally defined) are two different things, conceptually and empirically.

Distinguishing between rationality and intelligence helps explain how adolescents can be, at the same time, intelligent and irrational. In the mid 1990s one of us coined the term dysrationalia (an analogue of dyslexia) to refer to the existence of irrational thought and action in intelligent individuals (Stanovich, 1993). Dysrationalia occurs in adults as well as adolescents, and we tend to be surprised by it. We tend to be surprised when a physician loses all his pension funds in a speculative financial venture or when an educated professional ignores proven medical treatments and goes to Mexico for a quack therapy. Likewise, we are baffled when highly intelligent adolescents display poor decision making and subject themselves to such obvious risk factors as substance abuse, sexually transmitted diseases, and unsafe driving (Reyna & Farley, 2006).

In short, we find it paradoxical when smart people take disastrous actions. Readers will see in this chapter that this pattern should not be surprising—because there is nothing paradoxical at all in the dissociation between intelligence and rationality. Thus, the developmental trajectories of the latter must be studied in their own right, and this developmental work is summarized in the latter half of this chapter. Our emphasis is on the importance of the development of rational thinking in adolescence, which has been less well studied than other aspects of cognitive competence, such as intelligence.

THE MEANING OF RATIONALITY IN MODERN COGNITIVE SCIENCE

Many philosophers see rationality as one of the most important human values (Nathanson, 1994). As defined in modern cognitive science, rationality is important for a person’s well-being because it is a trait that helps us to achieve our goals. This idea is at odds with other characterizations that deem rationality either trivial (little more than the ability to solve textbook logic problems) or in fact antithetical to human fulfillment (as an impairment to
an enjoyable emotional life, for instance). However, these ideas about rationality derive from a restricted and mistaken view of rational thought—one not in accord with the study of rationality in modern cognitive science.

Dictionary definitions of rationality tend to be rather unspecific ("the state or quality of being in accord with reason"), and some critics who wish to downplay the importance of rationality have promulgated a caricature that restricts its definition to the ability to do the syllogistic reasoning problems encountered in Philosophy 101.

The meaning of rationality in modern cognitive science is, in contrast, much more robust and important. Cognitive scientists recognize two types of rationality: epistemic and instrumental. Epistemic rationality concerns how well beliefs map onto the facts of the world. Epistemic rationality is sometimes called theoretical rationality or evidential rationality (see Audi, 1993, 2001; Foley, 1987; Harman, 1995; Manktelow, 2004; Over, 2004).

The simplest definition of instrumental rationality is: Behaving in the world so that you get exactly what you most want, given the resources (physical and mental) available to you. Somewhat more technically, one could characterize instrumental rationality as the optimization of the individual’s goal fulfillment. Economists and cognitive scientists have refined the notion of optimization of goal fulfillment into the technical notion of expected utility. The model of rational judgment used by decision scientists is one in which a person chooses options based on which option has the largest expected utility¹ (see Baron, 2008; Dawes, 1998; Hastie & Dawes, 2001; Wu, Zhang, & Gonzalez, 2004).

One of the fundamental advances in the history of modern decision science was the demonstration that if people's preferences follow certain patterns (the so-called axioms of choice—things like transitivity [if A is preferred to B, and B is preferred to C, then A must be preferred to C] and freedom from certain kinds of context effects), then they are behaving as if they are maximizing utility: They are acting to get what they most want (Edwards, 1954; Jeffrey, 1983; Luce & Raiffa, 1957; Savage, 1954; von Neumann & Morgenstern, 1944). This is what makes people’s degrees of rationality measurable by the experimental methods of cognitive science. Although it is difficult to assess utility directly, it is much easier to assess whether one of the axioms of rational choice is being violated. This is why the seminal heuristics and biases

¹The principle of maximizing expected value says that the action that a rational person should choose is the one with the highest expected value. Expected value is calculated by taking the objective value of each outcome and multiplying it by the probability of that outcome and then summing those products over all of the possible outcomes. Symbolically, the formula is: Expected Value = Σ p_i v_i; where p_i is the probability of each outcome and v_i is the value of each outcome. The symbol Σ is the summation sign, and simply means "add up all of the terms that follow." The term utility refers to subjective value. Thus, the calculation of expected utility involves identical mathematics except that a subjective estimate of utility is substituted for the measure of objective value.
research program inaugurred by Kahneman and Tversky (1972, 1973; Tversky & Kahneman, 1974, 1983) focused on the causes of thinking errors.

Rationality has multiple components (Stanovich, 2011), and it is hard to measure the optimal functioning of all of these components. That is, it is hard to specify whether "perfect" rationality has been attained. Researchers have found it much easier to measure whether a particular rational stricture is being violated—that is, whether a person is committing a thinking error—rather than whether their thinking is as good as it can be. This is much like our judgments at a sporting event where, for example, it might be difficult to discern whether a quarterback has put the ball perfectly on the money, but it is not difficult at all to detect a bad throw.

In fact, in many other domains of life this is often the case. Performance errors are much easier to spot than is optimal performance. Essayist Neil Postman (1988) argued, for instance, that educators and other advocates of good thinking might adopt a stance more similar to that of physicians or attorneys. He pointed out that physicians find it hard to define "perfect health," but they are much better at the easier task of spotting disease, so they focus on the latter. Likewise, lawyers are much better at spotting injustice and lack of citizenship than defining "perfect justice" or ideal citizenship. Postman argued that like physicians and attorneys, educators might best focus on instances of poor thinking, which are much easier to identify, instead of trying to define ideal thinking. The literature on the psychology of rationality has followed this logic in that the empirical literature has focused on identifying thinking errors, just as physicians focus on disease. Cognitive biases imply some degree of irrationality, just as failure to display a cognitive bias becomes a measure of rational thought.

A substantial research literature, comprising hundreds of empirical studies conducted over several decades, has firmly established that people's responses sometimes deviate from the performance considered normative on many reasoning tasks. For example, people may assess probabilities incorrectly, test hypotheses inefficiently, violate the axioms of utility theory, or they do not properly calibrate degrees of belief; their choices may be affected by irrelevant context, they may ignore the alternative hypothesis when evaluating data, and they may display numerous other information processing

The term heuristics refers to why people sometimes make errors in choosing actions and in estimating probabilities—because they use mental shortcuts (heuristics) to solve many problems. Reliance on heuristics introduces systematic biases in actions and probability judgments, for example, when people rely on vivid examples rather than objective statistics to estimate probabilities. These biases can sometimes lead to errors in judgment (e.g., when a vivid stimulus is not very probable). In most instances, however, the use of heuristics is efficacious. The heuristics and biases research tradition focuses on the minority of instances where the use of heuristics leads to errors. It emphasizes that these situations may be a minority of the decisions we make in life, but that they can be extremely important ones (Milkman, Chugh, & Bazerman, 2009; Stanovich, 2009b).
biases (Baron, 2008; Evans, 2007; Gilovich, Griffin, & Kahneman, 2002; Kahneman & Tversky, 2000; Stanovich, 1999, 2004, 2009b). Demonstrating that descriptive accounts of human behavior sometimes diverge from normative models is a main theme of the heuristics and biases research program. The heuristic processes that lead to such errors are often evolutionarily adaptive (Cosmides & Tooby, 1992, 1996; Oaksford & Chater, 2007), but it is important to understand that the evolutionary adaptiveness of subprocesses does not imply that humans, as whole organisms, are rational (de Sousa, 2007; Hurley & Nudds, 2006; Samuels & Stich, 2004; Skyrms, 1996; Stanovich, 2004, 2011; Stanovich & West, 2003; Stich, 1990).

RATIONALITY AND INTELLIGENCE

Intelligence, as measured on many commonly used tests, is often separated into fluid and crystallized components, deriving from the Cattell/Horn/Carroll (CHC) theory of intelligence (Carroll, 1993; Cattell, 1963, 1998; Horn & Cattell, 1967). Sometimes termed the theory of fluid and crystallized intelligence (Gf/Gc theory), this theory posits that tests of mental ability tap, in addition to a general factor (g), a small number of broad factors, of which two are dominant (Geary, 2005; Horn & Noll, 1997; Taub & McGrew, 2004). Fluid intelligence (Gf) reflects reasoning abilities operating across a variety of domains—in particular, novel ones. It is measured by tasks of abstract reasoning such as figural analogies, Raven Matrices, and series completion. Crystallized intelligence (Gc) reflects declarative knowledge acquired from acculturated learning experiences. It is measured by vocabulary tasks, verbal comprehension, and general knowledge measures. Ackerman (1996) discussed how the two dominant factors in the CHC theory reflect a long history of considering two aspects of intelligence: intelligence-as-process (Gf) and intelligence-as-knowledge (Gc).

There is a large literature on the CHC theory and on the processing correlates of Gf and Gc (see Duncan et al., 2008; Geary, 2005; Gignac, 2005; Horn & Noll, 1997; Kane & Engle, 2002; McGrew & Woodcock, 2001; Taub & McGrew, 2004). In addition to Gf and Gc, other broad factors at the level termed stratum II are things like memory and learning, auditory perception, and processing speed (for a full account, see Carroll, 1993). However, most of these components are correlated with each other and with g, Gf, and/or Gc. When later in this chapter we discuss correlations between cognitive ability and rational thought, the cognitive ability indicators in the studies discussed will all be indicators of g, Gf, and/or Gc (Frey & Detterman, 2004).

In our theoretical approach to intelligence, our stance has been the standard one in cognitive science: to tie the concept to the actual empirical
operationalizations in the relevant literature. In the wider literature on intelligence in education, however, there is a contrary tradition. That contrary tradition leads to the distinction between broad and narrow theories of intelligence (Stanovich, 2009b). Broad theories include aspects of functioning that are captured by the vernacular term intelligence (e.g., adaptation to the environment, showing wisdom, displaying creativity) whether or not these aspects are actually measured by existing tests of intelligence. Narrow theories of intelligence, like that adopted in our discussion, confine the concept to the set of mental abilities actually tested on extant IQ tests.

In a sense, broad theories of intelligence are intellectually promiscuous in that they use a concept of intelligence that is not tied to existing tests of cognitive functioning. In fact, broad theories of intelligence have implicitly undervalued rational thinking by encompassing such skills under definitions of intelligence. The problem with such a stance is that it ignores the fact that none of the best-known intelligence tests measure rational thinking in any way. Our position is that broad theories of intelligence have been the source of much confusion both within the field and among the general public. They have led to overvaluing what IQ tests can reveal about intellectual functioning.

To think rationally means adopting appropriate goals, taking the appropriate action given one's goals and beliefs, and holding beliefs that are commensurate with available evidence. None of the currently used tests of intelligence assesses any of these functions (Perkins, 1995, 2002; Stanovich, 2002, 2009b; Sternberg, 2003, 2006). Thus, operationally, intelligence and rationality are separate things. However, although tests of intelligence fail to assess rational thinking directly, it could be argued that the processes that are tapped by IQ tests largely overlap with variation in rational thinking ability. If this were true, then measures of intelligence would correlate with measures of rational thinking despite the fact that the latter are not directly assessed on tests of the former. Variation in intelligence has been one of the most studied topics in psychology for many decades (Carroll, 1993; Deary, 2001; Geary, 2005; Lubinski, 2004), and the development of the cognitive abilities related to intelligence is likewise a central topic in developmental science (Anderson, 2005; Bjorklund, 2004; Kail, 2000). In contrast, variation in rational thought among adults has only recently been the focus of research (Bruine de Bruin, Parker, & Fischhoff, 2007; Stanovich & West, 1998c, 2000, 2008b).

Nonetheless, it has become apparent that intelligence and rational thinking skills are so modestly related that the former cannot be reliably used as a proxy for the latter. For example, some tasks in the heuristics and biases literature are related to intelligence, but the correlations are only in the range of .10–.40. For example, Stanovich and West (1997, 1998c, 1998d, 1999, 2000; see also Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Sá, West, & Stanovich, 1999; Toplak, Sorge, Benoît, West, & Stanovich, 2010;
Toplak & Stanovich, 2002; Toplak, West, & Stanovich, in press; West & Stanovich, 2003; West, Toplak, & Stanovich, 2008) found correlations with intelligence to be roughly (in absolute magnitude): .35–.45 for belief bias in syllogistic reasoning, .25–.35 for various probabilistic reasoning tasks, .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, .20–.30 with denominator neglect, and .05–.20 with various indices of Bayesian reasoning (see Stanovich, West, Toplak, 2011, for sources and examples of each of these paradigms). Other investigators have found relationships of a similar effect size between intelligence and a variety of tasks in the heuristics and biases literature (Bruine de Bruin, Parker, & Fischhoff, 2007; De Neys, 2006; DeShon et al., 1998; Handley et al., 2004; Klaczynski & Lavallee, 2005; Newstead et al., 2004; Parker & Fischhoff, 2005; Peters et al., 2006; Valentine, 1975).

The magnitude of these correlations may be underestimated, of course, because of unreliability of measurement or restriction of range. Our point, however, is not that they are near-zero correlations, or even small correlations, but only that they are so far below unity that a test of intelligence cannot be taken as a proxy for a test of rational thinking skill. For example, reading skill and intelligence can display correlations as high as .60, yet this magnitude of correlation leaves enough room for mismatches between the two that the discrepancies (e.g., low reading ability in the face of high performance on an IQ test) draw an extreme amount of attention from learning disability theorists.3 Our point with respect to the correlations between intelligence and measures of rational thought is similar. They may be higher than the raw correlations listed in the previous paragraph, but they will end up being estimated at so far below unity that we will not be able to use an IQ test score as an indirect measure of rational thinking. There will be a significant number of discrepancies (e.g., low rational thinking ability in the face of high performance on an IQ test) in any sample, and this discrepant pattern is of great theoretical and practical interest (Stanovich, 2009b, 2011).

Although some of these correlations are no doubt attenuated by restriction of range, there are other reasons to believe that the correlational values

---

3It is another question entirely whether educational policy should focus so much on this group of children. There is a substantial body of work indicating that the focus on discrepancy measurement in the domain of reading disability was a mistake. The proximal cause of most cases of reading difficulty—problems in phonological processing—is the same for individuals of high and low IQ (Stanovich, 2000; Stuebing, Barth, Molfese, Weiss, & Fletcher, 2009; Vellutino, Fletcher, Snowling, & Scanlon, 2004). Phonological processing is only modestly correlated with intelligence, so that cases of reading difficulty in the face of high IQ are in no way surprising and do not need a special explanation.
obtained in this research may be overestimated. In a commentary on this research on individual differences, Kahneman (2000) pointed out that the correlations observed may well have been inflated because most of the relevant studies used within-subjects designs that contain cues signaling the necessity of overriding automatic processing (Bartels, 2006; Fischhoff, Slovic, & Lichtenstein, 1979; Frisch, 1993; Kahneman & Tversky, 1982). Kahneman (2000) argued that between-subjects tests of the coherence of responses represent a much stricter criterion and perhaps a more appropriate one because “much of life resembles a between-subjects experiment” (p. 682).

Kahneman (2000; see also, Kahneman & Frederick, 2002) conjectured that these less transparent designs would reduce the observed relationships between intelligence and judgmental and decision-making tasks. Much less is known about the relation between cognitive ability and the tendency to make coherent judgments in between-subjects situations. However, in a series of studies, Stanovich and West (2007, 2008a, 2008b) attempted to examine a variety of effects from the heuristics and biases literature to see if intelligence was associated with these biases as they are displayed in between-subjects paradigms. In this series of experiments we found that, to a surprising degree, cognitive ability was independent of the tendency to show a variety of rational thinking biases. In university samples, intelligence was virtually independent of base rate neglect, framing effects, conjunction errors, outcome bias, anchoring effects, preference reversals, omission bias, and myside bias when these effects and biases were measured between subjects (for sources and examples of each of these paradigms, see Stanovich, West, & Toplak, 2011).

Thus, the judgment and decision-making tasks that represent the operational definition of rational thought are neither empirically nor conceptually identifiable with intelligence. Empirically, as just shown, the relationships are modest. Conceptually, we view rationality as a more encompassing construct than intelligence. To be rational, a person must have beliefs well calibrated with reality and must act appropriately on those beliefs to achieve goals. Additionally, in previous publications, we have argued that these aspects of rationality depend on certain thinking dispositions (see Stanovich, 2009a, 2009b, 2011; West, Toplak, & Stanovich, 2008).

No current test of intelligence or cognitive ability attempts to measure directly an aspect of epistemic or instrumental rationality. Neither does any test indirectly assess rational thought by examining any of the thinking dispositions that relate to rationality. For example, although intelligence tests do assess the ability to focus on an immediate goal in the face of distraction, they do not assess at all whether a person has the tendency to develop goals that are rational in the first place. Likewise, intelligence tests are good measures of how well a person can hold beliefs in short-term memory and manipulate those beliefs, but they do not assess at all whether a person has the
tendency to form beliefs rationally when presented with evidence. And again, similarly, intelligence tests are good measures of how efficiently a person processes information that has been provided, but they do not at all assess whether the person is a critical assessor of information as it is gathered in the natural environment.

Substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Many different studies involving thousands of subjects (e.g., Ackerman & Heggestad, 1997; Austin & Deary, 2002; Baron, 1982; Bates & Shieles, 2003; Cacioppo et al., 1996; Eysenck, 1994; Fleischhauer et al., 2010; Goff & Ackerman, 1992; Kanazawa, 2004; Kokis et al., 2002; Noffle & Robins, 2007; Zeidner & Matthews, 2000) 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 (e.g., conscientiousness, curiosity, diligence).

Other important evidence supports the conceptual distinction made here between intelligence and rational thinking dispositions. For example, across a variety of tasks from the heuristics and biases literature, it has consistently been found that rational thinking dispositions will predict variance in these 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 et al., 2002; Newstead et al., 2004; Macpherson & Stanovich, 2007; Parker & Fischhoff, 2005; Sá & Stanovich, 2001; Stanovich & West, 1997, 1998c, 2000; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007; Toplak & Stanovich, 2002; Toplak, West, & Stanovich, in press). Thinking disposition measures tell one about the individual's goals and epistemic values; they also index broad tendencies of pragmatic and epistemic self-regulation at a high level of cognitive control. The empirical studies cited indicate that these different types of cognitive predictors are tapping variance separable from intelligence.

**CATEGORIES OF ERROR IN JUDGMENT AND DECISION MAKING**

We have demonstrated in the previous section that an intelligence test is not a good proxy for individual differences in judgment and decision making skills. As a result, the developmental trajectories of the latter must be studied in their own right. We trace several such developmental trajectories in this chapter, but our review cannot be exhaustive. There are simply too many thinking errors and biases that have been demonstrated in the judgment and decision-making literature. In fact, a taxonomy of systematic error
types is badly needed, and several investigators have made some initial efforts at a classification scheme (e.g., Arkes, 1991; Oreg & Bayazit, 2009; Reyna, Lloyd, & Brainerd, 2003; Shah & Oppenheimer, 2008). We shall use a simplified and abbreviated version of our own taxonomy (Stanovich, 2009a, 2009b, 2011; Stanovich, Toplak, & West, 2008; Stanovich, West, & Toplak, 2011; Toplak et al., 2007) to organize our discussion of developmental trends. Our taxonomy is based around the finding that the human brain has two broad characteristics that make it less than rational. One is a processing problem and one is a content problem, and intelligence provides insufficient inoculation against both.

The processing problem is that we tend to be cognitive misers in our thinking. This has been a major theme throughout the past 40 years of research in the cognitive science of human judgment and decision making (Dawes, 1976; Simon, 1955, 1956; Taylor, 1981; Tversky & Kahneman, 1974). When approaching any problem, our brains have available various computational mechanisms for dealing with the situation. These mechanisms embody a tradeoff, however. The tradeoff is between power and expense. Some mechanisms have great computational power—they can solve a large number of problems and solve them with great accuracy. However, this power comes with a cost. These mechanisms take up a great deal of attention, tend to be slow, tend to interfere with other thoughts and actions we are carrying out, and require great concentration that is often experienced as aversive. In contrast, other brain mechanisms, often termed heuristic processes, are low in computational power but have the advantage that they are low in cost. These mechanisms do not permit fine-grained accuracy, but they are fast-acting, do not interfere with other ongoing cognition, require little concentration, and are not experienced as aversive.

We humans are cognitive misers in our basic tendency to default to heuristic processing mechanisms of low computational expense. Using less computational capacity for one task means that there is more left over for another task if they both must be completed simultaneously. This would seem to be adaptive. Nevertheless, this strong bias to default to the simplest cognitive mechanism—to be a cognitive miser—means that humans are often less than rational. Increasingly, in the modern world we are presented with decisions and problems that require more accurate responses than those generated by heuristic processing. Heuristic processes often provide a quick solution that is a first approximation to an optimal response. But modern technological societies are in fact hostile environments for people reliant on only the most easily computed automatic response. Think of the multimillion-dollar advertising industry that has been designed to exploit just this tendency. Modern society keeps proliferating situations in which shallow processing is not sufficient for maximizing personal happiness—precisely because many structures of market-
based societies have been designed explicitly to exploit the tendency toward heuristic processes. Being cognitive misers will seriously impede people from achieving their goals. Many effects in the heuristics and biases literature are the results of the human tendency to default to miserly processing: anchoring biases, framing effects, preference reversals, nondisjunctive reasoning, myside biases, and status quo biases, to name just a few (for a review and descriptions of these paradigms, see Stanovich, 2011; Stanovich et al., 2011).

The second broad reason why humans are less than rational represents a content problem. As described in the previous several paragraphs, rationality often requires that responses based on heuristic processing be overridden and replaced by responses that are more accurately computed. An aspect of human information processing that has been relatively neglected is that the override process is not simply procedural but instead utilizes content, that is, it uses declarative knowledge and strategic rules (linguistically coded strategies). In previous discussions, override has been treated as a somewhat disembodied process. The knowledge bases and strategies that are brought to bear have been given little attention. But in fact problems and gaps in these knowledge structures represent a second major class of reasoning error.

The term mindware was coined by Perkins (1995) to refer to the rules, knowledge, procedures, and strategies that a person can retrieve from memory to aid decision making and problem solving. If one is going to trump a heuristic response with conflicting information or a learned rule, one must have previously learned the information or the rule. If, in fact, the relevant mindware is not available because it has not been learned, the result is a thinking error that we have termed a mindware gap (Stanovich, 2009b; Stanovich et al., 2008). Mindware gaps can occur in a potentially large set of coherent knowledge bases in the domains of probabilistic reasoning, causal reasoning, logic, and scientific thinking (e.g., the importance of alternative hypotheses), the absence of which could result in irrational thought or behavior.

Mindware problems, in our view, are of two types: mindware gaps and contaminated mindware. Mindware encompasses a variety of declarative knowledge bases as well as some of the more content-laden strategies of rational thought, such as principles of scientific thinking and knowledge of probability. When an override of heuristic processing is necessary but the mindware for a substitute response is not available, then we have a case of a mindware gap. However, the subcategory of contaminated mindware is designed to draw attention to the fact that not all available mindware is helpful, either to goal attainment or to epistemic accuracy. In fact, some acquired mindware can be the direct cause of irrational actions that thwart our goals. Such effects thus define the subcategory of contaminated mindware. For example, one type of contaminated mindware that is discussed in the literature is mindware that contains evaluation-disabling properties (Blackmore,
1999; Dawkins, 1993; Distin, 2005; Dennett, 2006; Lynch, 1996; Stanovich, 2004). Some of the evaluation-disabling properties that help keep some mindware lodged in their hosts are: the promise of punishment if the mindware is questioned, the promise of rewards for unquestioning faith in the mindware, or the thwarting of evaluation attempts by rendering the mindware unfalsifiable.

THE DEVELOPMENTAL TRAJECTORIES OF RATIONAL THINKING SKILLS

The empirical literature on the development of rational thinking is still relatively sparse (see Byrnes, 1998; Jacobs & Klaczynski, 2005; Kokis et al., 2002; Reyna, Estrada, DeMarinis, Myers, Stanisz, & Mills, in press; Reyna, Lloyd, & Brainerd, 2003; Reyna & Rivers, 2008; Stanovich et al., 2008). More specifically, it might be said that the developmental literature is spread widely, but it is thin. We characterize this literature as thin because of the sheer number of different effects and biases that have been studied in the heuristics and biases adult literature (see Arkes, 1991; Baron, 2008; Evans, 1989, 2007, 2008; Gilovich, Griffin, & Kahneman, 2002; Johnson-Laird, 2006; Koehler & Harvey, 2004; Larrick, 2004; Lichtenstein & Slovic, 2006; McFadden, 1999; Nickerson, 2008; Reyna et al., 2003; Stanovich, 2009b, 2010). These tasks and effects represent the field of judgment and decision making in cognitive science. Continuous differences on these tasks and effects, as we previously discussed, define the construct of individual differences in rational thought. Although there have been some initial developmental studies on some of the tasks in the heuristics and biases literature, none of the tasks have been the subject of intense investigation. In this section, we discuss exemplar developmental studies involving many of the different tasks and biases of rational thinking.

A Problem of the Cognitive Miser: Vividness Effects

Affect substitution is a specific form of a more generic trick (attribute substitution) of the cognitive miser discussed by Kahneman and Frederick (2002): the substitution of an easy-to-evaluate characteristic for a harder one, even if the easier one is less accurate. One of the most common processing defaults of the cognitive miser is the tendency to default to vivid presentations of information and to avoid nonsalient presentations of evidence. For example, a picture would be vivid compared to a numerical presentation of evidence (Slovic & Peters, 2006).

In the heuristics and biases literature, a typical problem requires the participant 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 favors one of the bipolar decisions. The other evidence is a concrete case or vivid personal experience that points in the opposite direction. The classic Volvo versus Saab item (see Fong, Krantz, & Nisbett, 1986, p. 285) provides an example. In this problem, a couple is deciding to buy one of two otherwise equal cars. Consumer surveys, statistics on repair records, and polls of experts favor 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.

Kokis et al. (2002) adapted several problems such as this for children. One problem went as follows: Erica wants to go to a baseball game to try to catch a fly ball. She calls the main office and learns that almost all fly balls have been caught in section 43. Just before she chooses her seats, she learns that her friend Jimmy caught two fly balls last week sitting in section 10. Which section is most likely to give Erica the best chance to catch a fly ball?

(a) Definitely section 43  
(b) Probably section 43  
(c) Probably section 10  
(d) Definitely section 10

Selection of options a or b indicates the use of the aggregate base-rate information. Selection of options c or d indicates that the child is using the vivid information from a friend. Kokis et al. (2002) found a significant developmental trend whereby 13- to 14-year-olds displayed significantly less reliance on the vivid personal information than did a group of 10- to 11-year-olds.

Jacobs and Potenza (1991) found an analogous, and significant, developmental trend in the so-called object condition of their experiment where the problems were similar to those used by Kokis et al. (2002). However, in the so-called social condition of the Jacobs and Potenza study, the developmental trend was in the opposite direction—there was more reliance on the vivid information and less reliance on the more diagnostic statistical information by the older children. A consideration of the nature of the social problems reveals why this was the case. Here is an example of a social problem: In Juanita’s class, 10 girls are trying out to be cheerleaders and 20 are trying out for the band. Juanita is very popular and very pretty. She is always telling jokes and loves to be around people. Do you think Juanita is trying out to be a cheerleader or for the band? Here the statistical information points in the direction of band, but the personal information points in the direction of...
cheerleader. But to understand the diagnosticity of the indicant information in this problem, one must have knowledge of a social stereotype (that popular girls are drawn more to cheerleading than to band). Knowledge of this stereotype might well increase with age and thus be less available to the younger children. As a result, there is no conflict between the indicant and the base rate for the younger children.

Thus, the performance in the Jacobs and Potenza (1991) study is less inconsistent with the findings of Kokis et al. (2002) than may be apparent on the surface. The same is true of the social problems in a study by Davidson (1995). Finally, the developmental trend in Kokis et al. and in the object condition of Jacobs and Potenza (1991) is consistent with studies of individual differences of cognitive ability within an age group. Reliance on vivid individuating information is negatively correlated with cognitive ability (Kokis et al., 2002; Stanovich & West, 1998c, 1998d).

A Problem of the Cognitive Miser: Framing Effects

Framing effects represent the classic example of miserly information processing. For example, in discussing the mechanisms causing framing effects, Kahneman (2003) stated that “the basic principle of framing is the passive acceptance of the formulation given” (p. 703). The frame presented to the subject is taken as focal, and all subsequent thought derives from it rather than from alternative framings because the latter would require more thought. One of the most compelling framing demonstrations is from the early work of Tversky and Kahneman (1981).

Decision 1

Imagine that the United States is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows: If Program A is adopted, 200 people will be saved. If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. Which of the two programs would you favor, Program A or Program B?

Most people when given this problem prefer Program A—the one that saves 200 lives for sure. There is nothing wrong with this choice taken alone. However, inconsistent responses to another problem define a framing effect:

Decision 2

Imagine that the United States is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative pro-
grams to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows: If Program C is adopted, 400 people will die. If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die. Which of the two programs would you favor, Program C or Program D?

Most people when presented with Decision 2 prefer Program D. Thus, across the two problems, the most popular choices are Program A and Program D. The problem here is that Decision 1 and Decision 2 are really the same decision; they are merely redescriptions of the same situation. Program A and C are the same. That 400 will die in Program C implies that 200 will be saved—precisely the same number saved (200) in Program A. Likewise, the two-thirds chance that 600 will die in Program D is the same two-thirds chance that 600 will die (“no people will be saved”) in Program B. Many people show inconsistent preferences, their choice switches depending on the phrasing of the question. This is an example of a problem with very transparent equivalence. When presented with both versions of the problem together, most people agree that the problems are identical and that the alternative phrasing should not have made a difference.

Such a lack of so-called descriptive invariance is a very fundamental violation of some of the simplest strictures of rational thought (see Tversky & Kahneman, 1981, 1986). A theory of why these framing effects occur was presented in the prospect theory of Kahneman and Tversky (1979), which contains the key assumption that the utility function is steeper (in the negative direction) for losses than for gains. This explains why people tend to be more risk averse for gains than for losses (for an alternative explanation, see Kuhberger & Tanner, 2010; Reyna & Brainerd, 1991, 1995; Reyna et al., in press).

The literature on framing effects in adults is vast (see Epley, Mak, & Chen Idson, 2006; Kahneman & Tversky, 1984, 2000; Kuhberger, 1998; Levin, Gaeth, Schreiber, & Lauriola, 2002; Maule & Villejoubert, 2007; Schneider et al., 2005). However, the developmental literature is quite small. Obviously, the complexity of the problems has to be vastly reduced and made appropriate for children. Outcomes in developmental studies become small prizes that the children receive instead of the imaginary deaths or real money used in adult studies. Several investigators have creatively adapted framing paradigms for children, but the results of these experiments have not converged. Levin and colleagues (Levin & Hart, 2003; Levin, Hart, Weller, & Harshman, 2007) found no developmental trend for framing effects. Children (6- to 8-year-olds) were more risk averse for gains than for losses in the manner that prospect theory predicts, but the magnitude of the framing effects that they displayed was the same as that found for adults. Nonnormative
framing effects were found for 5-year-olds, 6-year-olds, and 9- to 10-year-olds in Schlottmann and Tring (2005), and in 14.8-year-olds in Chien, Lin, and Worthley (1996). In Schlottmann and Tring (2005), the framing effects were roughly similar in magnitude.

The results of these studies were not completely convergent with those of a study by Reyna and Ellis (1994; see also, Reyna, 1996). The data patterns in the Reyna and Ellis study were complex, however, and very variable over the ages studied. Reyna and Ellis predicted these results based on fuzzy trace theory, building on results with adults. Framing interacted with level of risk and magnitude of reward at certain ages. Briefly though, preschoolers' responses were consistent across frames: they displayed no framing effect. Second graders displayed a small reverse framing effect—they were more risk averse for losses. Fifth graders displayed a small reverse framing effect at the highest reward magnitude and the standard framing effect (more risk seeking for losses) only for lowest reward magnitude (see also Reyna & Farley, 2006).

The results of Levin and colleagues and those of Reyna and Ellis (1994) are consistent in one important respect, however: None of the studies reported evidence of a monotonically decreasing framing effect with age. This lack of developmental trend converges with studies of individual differences within an age group. Framing effects show very small (sometimes zero) correlations with cognitive ability, especially when tested in a between-subjects design (Bruine de Bruin et al., 2007; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999; 2008b; Toplak & Stanovich, 2002).

Override Failure: Denominator Neglect

As previously discussed, rationality often requires that responses based on heuristic processing be overridden and replaced by responses that are more accurately computed. But override is a capacity-demanding operation (see Stanovich, 1999, 2009b, 2011; Stanovich & West, 2000, 2008b). Thus, any tendencies toward miserly processing in a situation where override is required will result in a failure to substitute the superior response for the heuristic one. One of several phenomena in the heuristics and biases literature that illustrate the failure of heuristic override is the phenomenon of denominator neglect. Epstein and colleagues (Denes-Raj & Epstein, 1994; Kirkpatrick & Epstein, 1992; Pacini & Epstein, 1999; see also Reyna & Brainerd, 2008) demonstrated that it can result in a startling failure of rational judgment. Adults in several of his experiments were presented with two bowls that each contained clearly identified numbers of jelly beans. In the first were nine white jelly beans and one red jelly bean. In the second were 92 white jelly beans and eight red. A random draw was to be made from one of the two bowls, and if the red jelly bean was picked, the participant would receive a dollar. The participant could
choose which bowl to draw from. Although the two bowls clearly represent a 10% and an 8% chance of winning a dollar, a number of subjects chose the 100-bean bowl, thus reducing their chance of winning. The majority did pick the 10% bowl, but a healthy minority (from 30% to 40% of the participants) picked the 8% bowl. Although most of these participants in the minority were aware that the large bowl was statistically a worse bet, that bowl also contained more of the enticing winning beans—the eight red ones. In short, the tendency to respond to the absolute number of winners, for these participants, trumped the formal rule (pick the one with the best percentage of reds) that they knew was the better choice. That many subjects were aware of the poorer probability but failed to resist picking the large bowl is indicated by comments from some of them such as the following: “I picked the one with more red jelly beans because it looked like there were more ways to get a winner, even though I knew there were also more whites, and that the percents were against me” (Denes-Raj & Epstein, 1994, p. 823).

Kokis et al. (2002) found no significant trend for denominator neglect to decrease across their two age groups of 10- to 11-year-olds and 13- to 14-year-olds. However, Kokis et al. did find that cognitive ability was negatively correlated with denominator neglect to a significant degree. They used a paradigm very similar to that of Epstein and colleagues, but Klaczynski (2001b) altered the paradigm in an interesting way, by presenting options with equal probabilities and a response option that reflected the equivalent status of the two options. He had participants select from three alternatives: (a) a jar with one winning ticket out of 10, (b) a jar with 10 winning tickets out of 100, or (c) that it would not matter which jar was picked. Picking alternative b would indicate denominator neglect. Alternative c is the normatively correct response. Only a minority of participants chose the normatively correct response, but the normatively correct response did increase across three groups: early adolescents, middle adolescents, and young adults. In contrast, the majority response—denominator neglect—did not show a developmental trend. Amsel et al. (2008) also used Klaczynski’s paradigm and found no developmental increase in the normatively correct responding of 13-year-old middle school students and 23-year-old college students. Thus, even though the number of relevant developmental studies is quite limited, the existing research suggests that, at most, only minimal decline in denominator neglect occurs during adolescence.

Override Failure: Belief Bias

Clearer results have been obtained with another phenomenon caused by the failure of heuristic override: belief bias. Belief bias occurs when judgments of the believability of a conclusion interfere with judgments of logical
validity (Evans, Barston, & Pollard, 1983; Evans & Curtis-Holmes, 2005; Evans, Newstead, Allen, & Pollard, 1994). Consider the following syllogism:

Premise 1: All living things need water.
Premise 2: Roses need water.
Therefore: Roses are living things.
Judge the conclusion as either logically valid or invalid.

Premise 1 says that all living things need water, not that all things that need water are living things. So, just because roses need water, it does not follow from Premise 1 that they are living things. However, consider the following syllogism with exactly the same structure:

Premise 1: All insects need oxygen.
Premise 2: Mice need oxygen.
Therefore: Mice are insects.

Here it seems very easy to see that the syllogism is not valid. In both problems, prior knowledge about the nature of the world (that roses are living things and that mice are not insects) is becoming implicated in a type of judgment that is supposed to be independent of content: judgments of logical validity. In the rose problem, prior knowledge was interfering, and in the mice problem prior knowledge was facilitative.

Using syllogistic reasoning problems suitably modified for children, Kokis et al. (2002) found a significant developmental trend in which 13- to 14-year-olds displayed significantly less belief bias than did a group of 10- to 11-year-olds.

A much broader age span was examined by De Neys and Van Gelder (2009), who found a curvilinear developmental trend, whereby a group of 20-year-olds displayed less belief bias than either a group of 12-year-olds or a group of 65+-year-olds. Because fluid cognitive ability is likely to be on the decline for the latter subjects (Baltes, 1987; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002), the developmental trends in both Kokis et al. (2002) and De Neys and Van Gelder (2009) are consistent with studies of individual differences within an age group. There, it has been found that belief bias is negatively correlated with cognitive ability (De Neys, 2006; Handley et al., 2004; Sá et al., 1999; Stanovich & West, 1998c).

Override Failure: Prudently Discounting the Future

The tendency toward miserly processing is reflected in the choices often observed in delay discounting tasks, wherein an individual must choose between smaller-earlier and larger-later rewards. People are miserly because they do not expend the effort to inhibit the “gut-level” preference for the
earlier reward and wait for the larger, delayed reward. It takes cognitive effort to inhibit the preference for the earlier, smaller reward, and many people are miserly in not being willing to expend this effort.

*Delay discounting* refers to the decreasing value often placed on a future reward as the length of the delay needed to obtain the future reward increases. There is nothing inherently irrational about valuing an immediate reward more than a slightly greater future reward; indeed, it may sometimes be true that “a bird in the hand is worth two in the bush.” However, an extensive literature suggests that many people tend to imprudently discount the value they place on future rewards (e.g., Ainslie, 1975; Herrnstein, 1990; Kirby, 1997). One highly robust finding is that the rate at which rewards are discounted is nicely described by a highly concave function that is steeper for short delays than for long delays—called a *hyperbolic discount function* (Ainslie, 2001; Kirby, 1997).

Although almost everyone displays delay discounting that is consistent with such hyperbolic functions, there are large individual differences in the steepness of the initial portion of the curve. For some people, even relatively brief delays can result in drastic drops in the value placed on future rewards. These individuals may place little value on future rewards and, thus, may be particularly susceptible to making impulsive decisions (Ainslie, 1975, 2001). Kirby, Winston, and Santiesteban (2005) explored this possibility with college students. They estimated hyperbolic discount rates using a willingness-to-pay auction paradigm for real delayed monetary rewards (ranging from 1 to 43 days in delay and from $10 to $20 in monetary rewards). They found that steep discounting of future rewards was negatively correlated with college GPA, even after controlling for SAT scores. This finding suggests that students who steeply discount the value of future rewards may overvalue immediate alternatives to studying now, relative to the future rewards associated with obtaining higher grades later.

Adolescents, particularly young adolescents, are frequently characterized as being exceedingly focused on the present and inadequately valuing the future consequences of their present decisions. Steinberg et al. (2009) used a delay discounting task to explore this characterization with large, representative groups of 10- to 11-, 12- to 13-, 14- to 15-, 16- to 17-, 18- to 21-, 22- to 25-, and 26- to 30-year-olds. For the discounting task, participants indicated their preferences between immediate rewards of various monetary amounts and a delayed reward of $1,000. The length of the delays ranged from 1 day to 1 year. Although delay discounting was observed for each of the seven age groups in the study, participants in the two youngest age groups (ranging from 10 to 13 years old) had significantly higher discount rates than participants in the four oldest age groups (ranging from 16 to 30 years of age). Although the discounting rate for participants in the 14- to 15-year-old group was
between those of the younger and older groups, these differences failed to reach a level of significance. Thus, the amount of an immediate reward judged to be of equal value to a future reward of $1,000 was lower for participants 14 years of age and younger than for those 16 years of age and older. No significant developmental trends were found in the discount rates within this older age range. These findings are consistent with those of Green, Fry, and Myerson (1994; see also Green, Myerson, & Ostaszewski, 1999) with 12.1- and 20.3-year-olds. Interestingly, the developmental trend continued—with 67.9-year-olds showing the lowest discounting rate.

These developmental results are consistent with the studies that have examined cognitive ability correlations within an age group. A meta-analysis with 24 eligible studies by Shamosh and Gray (2008) reported that higher intelligence was negatively correlated with delay discounting (weighted r = −.23). The Steinberg et al. (2009) study discussed above reported a similar negative association between delay discounting and IQ (r = −.27). Shamosh et al. (2008) also reported that steeper discounting was associated with lower cognitive ability.

**Mindware Gaps: Probabilistic Reasoning**

As previously discussed, in addition to miser tendencies, the second category of cognitive errors is represented by mindware problems—both mindware gaps and contaminated mindware (see Stanovich, 2009b, 2011). Mindware gaps are common in the domain of probability knowledge and judgment (see the classic studies by Kahneman & Tversky, 1972, 1973). Conjunction effects represent rational thinking errors that arise because of mindware gaps in the domain of probability. Consider a problem that is famous in the literature of cognitive psychology, the so-called Linda problem (Tversky & Kahneman, 1983):

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

a. Linda is a teacher in an elementary school. ____

b. Linda works in a bookstore and takes yoga classes. ____

c. Linda is active in the feminist movement. ____

d. Linda is a psychiatric social worker. ____

e. Linda is a member of the League of Women Voters. ____

f. Linda is a bank teller. ____

g. Linda is an insurance salesperson. ____

h. Linda is a bank teller and is active in the feminist movement. ____
Most people make what is called a conjunction error on this problem. Because alternative (h) (Linda is a bank teller and is active in the feminist movement) is the conjunction of alternatives (c) and (f) the probability of (h) cannot be higher than that of either (c) (Linda is active in the feminist movement) or (f) (Linda is a bank teller). All feminist bank tellers are also bank tellers, so (h) cannot be more probable than (f)—yet many people in adult studies rate alternative (h) as more probable than (f), thus displaying a conjunction error.

Davidson (1995) reported the counterintuitive finding of a developmental trend of increasing conjunction errors with age. However, the stimuli in this study had the same problem as those in the social condition of the Jacobs and Potenza (1991) study—they depended on the knowledge of a stereotype that might increase with age. For example, one item went as follows:

Mrs. Hill is not in the best health and she has to wear glasses to see. Her hair is gray and she has wrinkles. She walks kind of hunched over. Do you think that Mrs. Hill is: a waitress in a local restaurant; an old person who has grandchildren and a waitress at a local restaurant; [and so on].

Knowledge of the stereotype might lead children to make a conjunction error, but the knowledge of the stereotype is undoubtedly more extensive among sixth graders (Davidson’s oldest group) than among second graders (Davidson’s youngest group).

Another way to state the problem with the Davidson (1995) stimuli is to say that the potential of the stimuli to trigger a representativeness judgment is confounded with age. Such stimuli are fine for certain questions, but they are not appropriate stimuli to use when addressing the generic question of whether the propensity to commit a conjunction error with age increases or decreases. Klaczynski (2001a) used stimuli that were less confounded in this manner and found no developmental trend in the number of conjunction errors from early adolescence (mean age, 12.4 years) to middle adolescence (16.3 years).

In contrast to Davidson’s findings, a number of studies using stimuli that were less confounded reported modest developmental trends of decreasing conjunction errors between early childhood and late childhood, and more substantial developmental trends of decreasing conjunction error between late childhood and young adulthood (Agnoli, 1991; Fischbein & Schnarch, 1997; Fisk, Bury, & Holden, 2006; Fisk & Slattery, 2005).

A developmental trend in favor of more optimal reasoning by older than younger adolescents has been found in other domains of probabilistic reasoning, including the gambler’s fallacy, the law of large numbers, and reliance on the representativeness heuristic for probabilistic judgment (Fischbein & Schnarch, 1997; Klaczynski, 2000, 2001a; Klaczynski & Narasimham, 1998).
For example, Fischbein and Schnarch (1997) examined a variety of probability problems in Israeli 10- to 11-, 12- to 13-, 14- to 15-, and 16- to 17-year-old students, and in college students. None of the students had had instruction in probability. In a “representativeness” problem, students had to decide whether someone who chose the lotto game sequence 1, 2, 3, 4, 5, 6 had a greater, a lesser, or the same chance of winning than someone who chose the sequence 39, 1, 17, 33, 8, 27. Nonconsecutive sequences in general are much more likely than are consecutive sequences. However, any specific sequence, whether consecutive or not, is no more likely than any other. Many students incorrectly chose the nonconsecutive sequence as being the more likely sequence, apparently because of its resemblance to—or intuitive representativeness of—lottery sequences in general. However, a clear developmental trend toward correctly responding was found.

The gambler’s fallacy was examined with a coin toss problem that asked the students about the chance of a coin coming up heads a fourth time after having come up heads on the three previous tosses (Fischbein & Schnarch, 1997). Once again, a developmental trend toward normative responding was found. Finally, Fischbein and Schnarch (1997) also examined the conjunction fallacy using a scenario that described a young man who appeared likely to be headed toward a career in medicine. The young man had registered at the university, and the question was whether he was more likely to be a student in the medical school or a student. The three younger groups of students were found to be much more likely to commit the conjunction error than were the two groups of older students. Taken together, these studies indicate a developmental trend of increasing probabilistic knowledge.

Contaminated Mindware: Pseudoscientific and Superstitious Beliefs

The other type of mindware problem that prevents rational thinking is caused by the presence of contaminated mindware. An important category of contaminated mindware is that of pseudoscientific beliefs. From a developmental standpoint, there is no a priori reason to expect many superstitious, paranormal, and pseudoscientific beliefs (e.g., beliefs in alien abduction, astrology, creationism, extrasensory perception, ghosts, quack medical or health claims) to be negatively associated with intellectual growth or intelligence, and many such beliefs undoubtedly are acquired as one ages. A case in point is a shocking claim made by Pat Robertson, elderly televangelist and host of the television’s The 700 Club, in the face of the horrific Haitian earthquake of 2010. Robertson’s claim was that Haiti had been cursed after a pact with the devil—that the people of Haiti had gotten together and sworn a pact with the devil in order to free them from French control (Parker, 2010).
Skepticism about superstitious, pseudoscientific, and paranormal beliefs can be acquired through science and other educational experiences. Preece and Baxter (2000) found that although the levels of superstitious and pseudoscientific beliefs were common, there was nonetheless a steady developmental trend toward increased skepticism about the superstitious and pseudoscientific beliefs in a large sample of 14- to 16-year-old students in England. Preece and Baxter speculated that science education may have contributed to this developmental trend, particularly because the level of skepticism continued to increase from 17- to 18-year-old science students to postgraduate pre-service science teachers. It is perhaps noteworthy, however, that Preece and Baxter intentionally avoided exploring beliefs with religious connotations. It is possible that these types of supernatural phenomena might be little affected by development.

Grimmer and White (1992) looked more directly at the role of specific educational training in a study of a large sample of Australian university students. Medical students showed the highest level of skepticism, science students showed the next highest level, and arts students were the least skeptical. Thus, these two studies reviewed above suggest that science education may play a role in reducing superstitious, pseudoscientific, and paranormal beliefs. However, the empirical data in this area are currently very limited, and considerably more research will be needed before anything can be confidently concluded about the role of development and educational experiences.

Multiply Determined Reasoning Problems

Many thinking errors in the heuristics and biases literature are multiply determined—that is, they result from a combination of miserly processing and mindware problems. One example of such a multiply determined reasoning problem is probably Wason's (1966) four-card selection task. In this task, the participant is told the following:

Each of the boxes below represents a card lying on a table. Each one of the cards has a letter on one side and a number on the other side. Here is a rule: If a card has a vowel on its letter side, then it has an even number on its number side. As you can see, two of the cards are letter-side up, and two of the cards are number-side up. Your task is to decide which card or cards must be turned over in order to find out whether the rule is true or false. Indicate which cards must be turned over.

The participant chooses from four cards labeled K, A, 8, 5 (corresponding to not-P, P, Q, and not-Q). The correct answer is to pick the A and the 5 (P and not-Q), but the most common answer is to pick the A and 8 (P and Q)—the so-called matching response.
In the four-card selection task, incorrect responses are probably due to a variety of cognitive errors (see Evans, 2007; Klauer, Stahl, & Erdfelder, 2007; Osman & Laming, 2001), some of them reflecting miserly processing and some of them mindware gaps. The much-discussed matching bias (unreflectively picking the elements mentioned in the rule) evident in the task (Evans, 1972, 1998; Evans & Lynch, 1973) is clearly a case of miserly processing. However, it is perhaps also the case that the matching response goes uncorrected because of insufficiently instantiated mindware that embodies the importance of thinking about alternative hypotheses. Developmental and individual-differences studies of the task are consistent, however. Superior performance on the task is associated with development (Klaczynski, 2001a; Overton, Byrnes, & O’Brien, 1985) and with cognitive ability (DeShon, Smith, Chan, & Schmitt, 1998; Stanovich & West, 1998a; Toplak & Stanovich, 2002; Valentine, 1975).

As another example of a multiply determined cognitive bias, consider myside bias. Critical thinking is often thought to entail the ability to decouple prior beliefs and opinions from the evaluation of evidence and arguments (Baltes & Staudinger, 2000; Evans, 2002; Kuhn, 2001; Johnson-Laird, 2006; Perkins, 1995, 2002; Sternberg, 2003). People display myside bias when they evaluate evidence, generate evidence, and test hypotheses in a manner biased toward their own opinions (Greenhoot, Semb, Colombo, & Schreiber, 2004; Klaczynski & Lavallee, 2005; Nussbaum & Kardas, 2005; Perkins, 1995; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a; Toplak & Stanovich, 2003; Wolfe & Britt, 2008). Myside bias surely derives in part from the miserly tendency to build the easiest cognitive model: a model from a single perspective—one’s own. Contributing additionally, however, might be contaminated mindware that structures our knowledge of the world from an egocentric perspective.

Developmental trends in myside processing have been studied by Klaczynski and colleagues (1997; Klaczynski & Gordon, 1996; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Narasimham, 1998). They presented participants with flawed hypothetical experiments that led to either opinion-consistent or opinion-inconsistent conclusions and evaluated the quality of the reasoning used when the participants critiqued the flaws in the experiments. In these experiments, myside bias effects were also evident; participants found more flaws when the experiment’s conclusions were inconsistent than when they were consistent with their opinions and beliefs. However, there was no developmental trend for myside bias to increase or decrease—at least for the early, middle, and late adolescent groups that were the focus of most the Klaczynski lab’s research.

The lack of a developmental trend in myside bias in the Klaczynski group’s studies is consistent with studies of individual differences within an
age group. Across a variety of myside paradigms, there has been little evidence that myside bias is associated with cognitive ability (Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Macpherson & Stanovich, 2007; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a; Toplak & Stanovich, 2003).

**OVERALL DEVELOPMENT TRENDS IN RATIONAL THOUGHT**

Collectively, the biases discussed in this chapter (and many others in the heuristics and biases literature not covered here) define departures from rationality and hence indirectly index rationality itself. The minimal conclusion to be drawn from the body of developmental work taken as a whole is that children unequivocally do show the biases that have been displayed in the adult literature: vividness effects, framing effects, denominator neglect, belief bias, conjunction errors, myside bias, and so forth.

Beyond this minimal conclusion, however, we must be tentative because of the rather sparse nature of the empirical literature. Very few of the areas covered in this review have been the subject of enough research to establish developmental trends with confidence. This caveat aside, the suggestive trends that are in the literature vary considerably across the various cognitive biases. Across the age ranges studied, there appear to be developmental increases in the avoidance of belief bias, analytic responding in the selection task, probabilistic reasoning, and reliance on statistical information in the face of conflicting personal testimonials. All of these trends are in the direction of increasingly rational thought, at least according to accepted normative standards.

These findings contrast, however, with the trends apparent in developmental studies of myside bias and framing effects. Neither of these biases was attenuated by development (see other such effects in Table 3 of Reyna & Farley, 2006). The lack of developmental decreases in these two biases is interestingly convergent with findings that neither bias displays much of a correlation with intelligence (Stanovich & West, 2007, 2008b). These aspects of rational thought are thus quite independent of cognitive ability, reinforcing the fundamental point with which we opened this chapter: Measures of cognitive ability, as traditionally defined, fail to assess degrees of rationality.

In the case of conjunction errors and contaminated mindware, the literature on developmental trends is simply too inconsistent and sparse to warrant any strong conclusions at this point. One reason that developmental trends in some of these areas might be unclear is that whether or not 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 development, 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 simulation of an alternative response by the analytic system. 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 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 in which fluent adult performance looks very heuristic (Brainerd & Reyna, 2001; Ericsson & Charness, 1994; Kahneman & Klein, 2009; Klein, 1998; Reyna, Lloyd, & Brainerd, 2003; Reyna et al., 2005; Shiffrin & Schneider, 1977).

The difficulty of pinning down some of these developmental trends highlights how thin the literature is on children’s rational thought. One cannot amalgamate all of the studies reviewed here (as well as others not reviewed) as evidence on a single issue; doing so would misleadingly suggest that the literature is more extensive than it is. Rational thought spans many domains. It encompasses many different thinking dispositions and knowledge domains, each of which has been investigated separately in the adult literature. For example, there is an enormous adult literature on conjunction effects (Fisk, 2004), on framing (Kahneman & Tversky, 2000), on base rate use (Koehler, 1996), and on every one of the myriad effects in the heuristics and biases literature. A parallel effort in each of these domains will be necessary in order to fully understand the development of rational thought.

Understanding the development of rationality will clearly be a tall order. It will be worth the effort, however, not just for scientific reasons. Assumptions about the nature and development of rationality are implicated in judgments of legal responsibility. Reyna and Farley (2006) emphasized how background assumptions about adolescent rationality frame efforts to change adolescent risk behavior. For example, theories that stress adolescent feelings of invulnerability serve to avoid the attribution of irrationality to adolescents who engage in high-risk behavior. If these adolescents have strong feelings of invulnerability, or if they overestimate the probabilities of negative outcomes but weight rewards sufficiently highly, then a consequentialist calculation might well make engaging in high-risk behaviors rational for them.
An alternative approach, one supported by the research in the heuristics and biases tradition, would relax the rationality assumption and conclude instead that some of these adolescent behaviors violate standard rational strictures. Such a stance finds additional motivation from an observation that Reyna and Farley (2006) discussed: Many adolescents making poor choices are alienated from the choices they make. As Reyna and Farley put it, "People who take unhealthy risks often agree that their behavior is irrational, on sober reflection, but they gave in to temptation or were not thinking at the time of the decision, and are worse off for having done so" (p. 35). Instead of the dated economics-like assumption of adolescents as coherent rational actors, modern approaches in decision science highlight the image of a decision maker in conflict (Evans, 2008, 2009; Stanovich, 2004, 2009b, 2011). This comports well with the fact that many adolescents with behavior problems will indeed verbally reject their own behavior. Such a philosophical reorientation could, as Reyna and Farley (2006) demonstrated, have profound implications for how we interpret many findings in the area of adolescent decision making.

REFERENCES


JUDGMENT AND DECISION MAKING IN ADOLESCENCE 367


JUDGMENT AND DECISION MAKING IN ADOLESCENCE 373


JUDGMENT AND DECISION MAKING IN ADOLESCENCE

375


