Child Development, May/June 2001, Volume 72, Number 3, Pages 844–861

Analytic and Heuristic Processing Influences on Adolescent Reasoning and Decision-Making

Paul A. Klaczynski

The normative/descriptive gap is the discrepancy between actual reasoning and traditional standards for reasoning. The relationship between age and the normative/descriptive gap was examined by presenting adolescents with a battery of reasoning and decision-making tasks. Middle adolescents ($N = 76$) performed closer to normative ideals than early adolescents ($N = 66$), although the normative/descriptive gap was large for both groups. Correlational analyses revealed that (1) normative responses correlated positively with each other, (2) nonnormative responses were positively interrelated, and (3) normative and nonnormative responses were largely independent. Factor analyses suggested that performance was based on two processing systems. The "analytic" system operates on "decontextualized" task representations and underlies conscious, computational reasoning. The "heuristic" system operates on "contextualized," content-laden representations and produces "cognitively cheap" responses that sometimes conflict with traditional norms. Analytic processing was more clearly linked to age and to intelligence than heuristic processing. Implications for cognitive development, the competence/performance issue, and rationality are discussed.

INTRODUCTION

In recent years, both basic and applied researchers have become increasingly interested in discerning the mechanisms underlying adolescent decision-making (Byrnes, 1998). Cognitive research on decision-making has concentrated on the development of abilities that are frequently associated with intellectual maturity and on the correlations between these abilities and aspects of decision-making (e.g., awareness of costs and benefits). The unfortunate outcome of this exclusive focus on higher order competencies (e.g., formal operations) has been a misleading picture of the cognitive foundations on which decision making rests.

This allegation is based on findings from the "heuristics and biases" research program (see Kahneman, Slovic, & Tversky, 1982) and the controversies surrounding those findings (e.g., Gigerenzer, 1996; Kahneman & Tversky, 1996). Specifically, the long-held assumption that logical, computational processing is essential to rational decision-making has been taken to task. Recent theoretical advances emphasize that judgment and decision-making (JDM) is heavily influenced by preconscious processes (Epstein, 1994; Evans & Over, 1996). Advocates of these "two-process" theories subscribe to the belief that rationality depends on both conscious, "analytic" processing and preconscious, "heuristic" processing.

In two-process theories, cognition is seen as developing along two dissociated trajectories—one directed toward increases in computational processing and the capacity to decontextualize reasoning from problem content; the second directed toward heuristic, highly contextualized processing (Stanovich, 1999). The emergence of two-process theories has thus posed serious challenges to views of development as a unidirectional progression from intuitive processing to predominantly logico-mathematical processing (e.g., Piaget & Inhelder, 1951/1975).

The research presented herein examined adolescent decision-making from a two-process perspective. Early and middle adolescents were presented with a battery of reasoning and JDM tasks. Age differences in these tasks, intertask correlations, and the relations between task performance and an index of general ability were examined. Before further specifying the goals of this research, findings that have led to the current popularity of two-process theories, and basic tenets of these theories, are reviewed.

Empirical Evidence for Two Cognitive Systems

Data accumulated over the past 3 decades indicate the inadequacy of information processing and Piagetian theories to explain the flexibility and variability of children's responses to numerous cognitive tasks. Much of this evidence has not only falsified predictions generated from these traditions, but has also shown that satisfactory accounts of intellectual growth must explain the perplexing frequency of errors on simple logical problems (e.g., Wason, 1966), the relationships between reasoning and memory (Reyna &

Brainerd, 1995), and the observation that variability, rather than consistency, is the norm in everyday cognition and its development (Siegler, 1996). Although this evidence has numerous origins, only three of these are focused on for the sake of brevity.

First, age-related increases in certain nonnormative response tendencies are difficult to explain without reference to two cognitive systems. Jacobs and Potenza (1991), for instance, found that reliance on statistical evidence on asocial decision tasks (e.g., about bicycles) increased with age. On logically isomorphic social problems, however, the opposite trend was observed: With increasing age, children relied more on the "representativeness heuristic" (i.e., the extent to which individual cases conform to existing schemata) and less on statistical evidence. The tendency for older children to commit the "conjunction fallacy," that is, to judge $p(AB)$ as more probable than $p(A)$ or $p(B)$, more than younger children has also been attributed to increased reliance on representativeness and less on statistical evidence (Davidson, 1995). Here, normative responses refer to those historically advocated by logicians, philosophers, and decision theorists. Nonnormative responses include a variety of responses—such as the heuristics and biases described by Kahneman, Tversky, and others (see Evans, 1989; Kahneman et al., 1982)—that have no basis in formal theories of reasoning (e.g., Inhelder & Piaget, 1958; Rips, 1994) and errors that arise from faulty logical processing.

The terms "normative" and "nonnormative" are used in this historical sense rather than as synonyms for "correct" and "incorrect" because arguments have been made that, under some conditions, purportedly nonnormative responses are adaptive (see Evans & Over, 1996). Although terms such as "computational" and "noncomputational" might be used instead, arguments have been made that some of the normative responses studied here are not products of logical analysis. Thus, the normative–nonnormative distinction was retained, but with the caveats that (1) the distinction is a loose one and is not absolute, (2) if traditional standards for correctness are applied, some normative responses fall short of these standards (these responses are discussed in the Method section), and (3) responses traditionally termed normative are sometimes maladaptive, and responses traditionally termed nonnormative are sometimes adaptive (see Evans & Over, 1996; Jacobs & Narloch, 2001; Kahneman et al., 1982; Klaczynski, 2001; Reyna et al., in press).

Similar evidence can be found across disparate methodological paradigms and dimensions of cognitive development. Despite knowledge of normative computational strategies, age (under certain conditions) is positively associated with (1) making probability judgments based on simple, cognitively economical, strategies (e.g., ignoring denominators in ratio problems; Brainerd, 1981), (2) changing decisions as a function of the "framing" of logically identical problems (Reyna & Ellis, 1994), (3) making nonlogical "transitive" inferences regarding social relationships (e.g., "A is a friend of B. B is a friend of C. Therefore, A and C are friends"; Markovits & Dumas, 1999), (4) committing deductive reasoning fallacies (Klaczynski & Narasimham, 1998a), and (5) imputing false beliefs to others (Mitchell, Robinson, Isaacs, & Nye, 1996). Because they are systematic and yet violate formal rules of inference, these developmental trends must arise from a cognitive system that does not rely on logico-mathematical processing.

A second source of evidence is available in the memory literature. For example, children and adults sometimes "remember" information without conscious awareness that they are remembering. When incomplete items are presented during an early phase of an experiment, both children and adults later identify the complete items more accurately and quickly than nonprimed items—despite being unaware that they had previously examined the primed items (Bargh & Chartland, 1999; Hayes & Hennessy, 1996). Other evidence similarly suggests separate systems for conscious and unconscious memory. Newcombe and Fox (1994), for example, found that few 9- to 10-year-olds accurately recognized pictures of preschool classmates. Skin conductance changes for most children, however, were greater for the pictures of actual classmates than for "false" classmates. Recognition was unrelated to performance on the physiological measure, thus indicating conscious–unconscious memory independence (Lie & Newcombe, 1999). This conclusion is reinforced by findings that, unlike explicit memory, in implicit memory age differences are minimal (Schneider & Bjorklund, 1998).

Implicit memory research, and research on other forms of unintentional memory, indicates that parallel processing allows experiences to be encoded at several levels. Generally speaking, a given experience is represented by verbatim traces (which entail correspondence to problem details) and by gist traces (i.e., holistic abstractions of patterns). Of critical importance are data showing that development is marked by increased reliance on gist representations (Brainerd & Gordon, 1994). This verbatim→gist shift has clear adaptive value: Compared with verbatim traces, gist is less susceptible to interference and forgetting, is more cognitively economical, and, in part because it is less cumbersome, lends itself more easily to higher order cognitive operations (Reyna & Brainerd, 1995). Available evidence thus implies that the two memory systems encode different representational contents and follow distinct developmental paths.

Finally, findings that verbatim memory and reasoning are sometimes independent violate the Piagetian assumption that memory is reconstructive (and thus arises from reasoning) and the information-processing assumption that memory is necessary for reasoning. In an influential series of investigations, Brainerd and Kingma (1984, 1985) showed that memory for premise information on several Piagetian tasks (e.g., transitive inference, class inclusion) was unrelated to reasoning. For example, children who remember precise quantitative information (e.g., 5 dogs, 9 horses, 11 cows) are no more or less likely than children who incorrectly remember such details to commit inclusion errors (e.g., there are more cows than animals).

Operations performed on the "deep" (i.e., logical) structure of tasks thus depend on representations that are functionally dissociated from the processes responsible for encoding surface (i.e., verbatim) information. One outcome of this dissociation is that, in certain task environments, explicit reasoning and verbatim memory operate independently. This finding, together with research described earlier, necessitates conceptualizations of cognitive development that emphasize separate roles for verbatim and gist representations, parallel processing at the conscious and preconscious levels, and different developmental trajectories for inferences that arise from conscious, "deep" processing and those that arise from "peripheral" processing.

Properties of Heuristic and Analytic Processing

The foregoing discussion highlights several features of two-process approaches, such as the assertions that decision-making is jointly determined by interactions between two cognitive systems (Epstein, 1994) and that preconsciously extracted representations often form the basis for consciously made decisions (Evans, 1996). The former claim requires the caveat that task characteristics (e.g., familiarity), context (e.g., social demands for accuracy), and individual difference variables (e.g., epistemic beliefs, intelligence) interact to determine which processing system is predominant on a given task (Stanovich, 1999). The latter claim is critical in that representations are also essential determinants of the processing system accessed in a given situation.

The properties and characteristics of the two systems differ at several levels. Heuristic system processing is relatively rapid, enables automatic recognition of environmental features (e.g., facial cues), and facilitates information mapping onto and assimilation into existing knowledge categories. Relative to analytic processing, heuristic processing occurs at the periphery of awareness, requires little cognitive effort, and thus frees attentional resources for computationally complex reasoning.

When heuristic processing is predominant on a task, responses have no basis in reasoning in the "usual" sense; that is, computational analyses and attempts to break problems down into discrete components are absent; little or no attention is paid to formal rules of inference or decision-making. For instance, although heuristics appear to derive from well-learned, automated rules, such rules are applied "thoughtlessly" (i.e., without concern for their limitations; see Arkes & Ayton, 1999). Heuristic processing also may predominate when tasks activate stereotypes, personal "theories" (e.g., of personalities), and vivid or salient memories (Kahneman & Tversky, 1972; Klaczynski, 2000). Phenomenologically, judgments arrived at heuristically feel intuitively correct, but the basis for this feeling is often difficult to articulate (Epstein, 1994).

Analytic processing is consciously controlled, effortful, and deliberate. Successful analytic reasoning depends on the acquisition of abilities that are frequently prescribed as normative for reasoning and decision-making (Epstein, 1994). Analytic competencies include the higher order abilities that enable reasoning consistent with the rules of formal logic, decisions based on comparisons between a priori probabilities, and accurate calibration of one's abilities. Unlike heuristic processing, analytic processing is directed toward precise inferences.

The two systems are assumed to operate on different task representations. Heuristic processing generally operates on "contextualized" representations that are heavily dependent on problem content (e.g., familiarity) and semantic memory structures (e.g., stereotypes). Analytic processing operates on "decontextualized" representations in which the underlying structure of a task is decoupled from superficial content and which thereby facilitate logico-computational operations (Stanovich & West, 1997).

The representation–processing system relation is considerably more complex than portrayed here. For example, decontextualized representations increase the probability of analytic processing but do not guarantee such processing. Even if analytic processing is engaged, normative solutions are not ensured because representations may be misleading, inappropriate reasoning principles may be applied, appropriate principles may be misapplied, or heuristic processing may interfere with reasoning despite conscious attempts to reason analytically (Chen & Chaiken, 1999; Klaczynski & Narasimham, 1998b). Contrasting positions and detailed explications of this relation are available in Reyna and Brainerd (1995), Reyna et al. (in press), Stanovich (1999), Evans and Over (1996), Klaczynski (2001), and Muller, Sokol, and Overton (1998).

The Present Investigation

Some of the tasks used in this study were selected because they are thought to bring analytic and heuristic processing into conflict (Denes-Raj & Epstein, 1994; Stanovich & West, 1998a). Such tasks are optimal for studying two-process predictions because the two systems "pull" for different solutions. Tasks were also selected because they involve analytic competencies usually acquired by the age of 14, but normative responding on several of the tasks is surprisingly poor. Unfortunately, most researchers have focused on individual task performance and on college students. Considerable information can be gained, however, by examining patterns of covariation among tasks and relations to psychometric intelligence and age (Stanovich, 1999).

Consider arguments that nonnormative responses (NNRs) reflect random performance errors or "momentary lapses of reasoning" (Stein, 1996); theorists espousing this view would be hard-pressed to explain systematic covariation among such responses. Similarly, significant correlations among NNRs render less plausible the possibility that some NNRs result from inaccurate representations, others result from misapplication of normative principles, and still others result from heuristic processing. Instead, positive manifold may indicate that predominantly heuristic processing, activated by highly contextualized representations, is their shared origin.

Further evidence that certain NNRs are produced heuristically would be gained if they were unrelated to psychometric intelligence. This is because different types of intelligence are believed to underlie heuristic and analytic processing (Epstein, 1994); traditional assessment instruments are intended to measure only analytic intelligence. If NNRs are merely analytic processing failures, they should correlate negatively with intellectual ability.

The argument for normatively correct responses (NCRs) to different tasks is similar: Positive manifold would imply a consolidated analytic system. Because measures of psychometric intelligence are relatively "pure" indices of analytic competence, significant correlations between NCRs and general ability would further support this hypothesis. If analytic and heuristic processes "pull" for different solutions, then only solutions associated with strong analytic pulls should correlate with intelligence.

Finally, if analytic competencies develop rapidly during the early years of adolescence (Ward & Overton, 1990), then NCRs should be made more often by middle adolescents (i.e., older than 14 years of age) than by early adolescents. Developmental expectations for nonnormative responses were not as clear. Research reviewed earlier indicates developmental increases in various NNRs during childhood. During adolescence, however, some NNRs appear to increase (Klaczynski & Narasimham, 1998b), but others seem to remain stable (Klaczynski, 2001).

In sum, developmental differences in responding to several JDM tasks were explored with the expectation that NCRs would increase with age. Correlations among NCRs and among NNRs, and with an index of psychometric intelligence, were examined to test the hypothesis that different processing systems predominate when these responses are produced.

METHOD

Participants

Sixty-six early adolescents (30 male, 36 female, M = 12.4 years, SD = 1.78) and 76 middle adolescents (38 male, 38 female, M = 16.3 years; SD = 1.23) participated in this study. The early adolescents were drawn from the seventh and eighth grades of a public elementary school. The middle adolescents were tenth- and eleventh-grade students enrolled in a public high school. Both schools service families from the lower middle to middle socioeconomic classes.

Procedure

Experimental sessions were conducted in rooms at participants' schools in two 35 to 45 min sessions that were separated by 1 day. Two sets of problems (half on Day 1, half on Day 2) were administered to groups of 3 to 8 adolescents. Problems (with the exceptions noted subsequently) were presented in one of four random orders determined before data collection.

Verbal Ability

Before problem presentation, the Primary Mental Abilities (PMA) Verbal Meaning test (Thurstone, 1962) was administered. This test was selected because vocabulary is the best single predictor of global intelligence scores and because the construct validity of this test has been demonstrated in other studies of adolescent reasoning (Klaczynski, 2000). For each of 64 items, the task is to select which of four words has the same meaning as a target word. Participants were given 5 min to complete as many items as possible.

Reasoning and Decision-Making Tasks

Eight tasks from the heuristics and biases literature were used. The tasks were three conditional reasoning problems, two contingency detection problems, a reasoning calibration item for each conditional reasoning and contingency problem, three statistical reasoning problems, four conjunction problems, two "gambler's fallacy" problems, two outcome bias problems, and a hindsight bias problem. Examples of each are presented in the appendix.

To maintain consistency with previous research, the response format for each task was similar to the format typically found in the adult literature. For most tasks, a distinction between normative and nonnormative responses was possible (see, however, the comments on normative responses on page 845). For the outcome bias and hindsight bias problems, however, the normative–nonnormative distinction was more ambiguous because responses were on continuous scales. For ease of presentation, and because they correlated positively with other NNRs, responses to the outcome bias and hindsight bias problems were labeled as nonnormative.

1. Conditional reasoning. In Wason's (1966) selection task, an "if *p* then *q*" conditional rule is presented. The rule is followed by four alternatives, representing the affirmation of the antecedent (*p*), the affirmation of the consequent (*q*), the denial of the antecedent (*not p*), and the denial of the consequent (*not q*). Each alternative (card) has an instance of the antecedent or its denial on one side and an instance of the consequent or its denial on the other side. The task is to select those cards that would allow the truth status of the rule to be tested with certainty. The normative (falsification) response is to select the *p* and the *not q* cards because only information on the back sides of these cards can falsify the rule.

Of the three conditional rules, two (one familiar, one unfamiliar) assessed reasoning about deontic relationships, that is, prescriptive rules regarding what should or ought to be done (Manktelow, Sutherland, & Over, 1995). Note that some theorists have argued that deontic selection task—from here on, DST—problems do not create analytic versus heuristic conflicts. Analyses were collapsed over familiarity because falsification solutions on these problems correlated highly, $r = .64$, $p < .001$, and were related similarly to age.

The third selection task problem contained an arbitrary rule with no meaningful association between the antecedent and the consequent (from here on, AST). Typically, fewer than 10% of participants provide the falsification response to such problems. Instead, the modal response is selection of the *p* and *q* cards. Evans (1989, 1996) provides evidence that the *pq* pattern reflects a matching bias in which selections are based on heuristics to select cards whose values correspond to those in the rule. On the basis of this and other evidence, Evans (1996; Evans & Over, 1996) asserts that the *pq* response should be considered normative for AST problems, and argues that the arbitrary task "may fail to elicit any cognitive processes of the type that we would wish to describe as 'reasoning'" (1996, p. 224).

Questions over the normative status of *pq* and falsification responses on AST have been fueled recently by the findings of Stanovich and West (1998c). In their research, falsification responses and "*p* card only" responses were associated with higher intellectual ability than *pq* responses. One construal of this finding is that the analytic system is predominant for falsification *and* for "*p* only" responses. Despite analytic predominance, however, those who make "*p* only" selections are unable to generate complete falsification answers. Under the assumption that the analytic system is predominant both for complete falsification solutions and for "*p* only" solutions, these two response patterns were collapsed into a single arbitrary analytic score. (This decision was made, in part, because complete falsification responses on the AST were uncommon; see Table 1.)

Because different processes may underlie reasoning on the DST and AST problems, they were analyzed separately. Total deontic falsification and deontic *pq* scores could range from 0 to 2. Arbitrary falsification, arbitrary *pq*, and arbitrary analytic scores could range from 0 to 1.

2. Contingency detection. These two problems were based on those created by Wasserman, Dorner, and Kao (1990). In each problem, a brief scenario described an investigation of the relationship between two variables (e.g., between a drug used to combat a disease and drug efficacy). A 2 (cause: A or B) × 2 (effect: present or absent) contingency table showing four outcomes was then presented. Participants rated the effectiveness of the putative cause on a 5-point scale (1 = had a very negative effect; 2 = had a somewhat negative effect; 3 = had no effect; 4 = had a somewhat positive effect; 5 = had a very positive effect). In one problem (see the Appendix), the correlation between cause and effect was 0 (therefore, 3 was the correct response). In the second problem, there was a moderate negative correlation between efficacy of a new teaching method and grades: The grades of 8 of 15 students (53%) who received the new method declined; of 30 students in the traditional method class, the grades of 9 students (30%) declined. Responses were considered correct when participants indicated that the new method had "a somewhat negative effect" or "a very negative effect." On each problem, correct responses depended on comparing ratios rather than focusing on absolute numbers in the putative cause A-effect present (e.g., new method: 8 students with lower grades) and cause B-effect present (e.g., traditional method: 9 students with lower grades) cells. This phenomenon, known as denominator neglect (Reyna et al., in press) and ratio bias (Epstein & Pacini, 1999), is common among both adults and adolescents.

Normative contingency scores, which could range from 0 to 2, provided rough indicators of the tendency to take denominators into account when making covariation judgments. In the first problem, sample sizes were small; hence, respondents who considered both covariation information and sample size would have been hard pressed to draw firm conclusions. In the second problem, sample sizes were larger, but judgments of the negative impact (i.e., "somewhat" or "very") of the new teaching method clearly involved a degree of subjectivity. The validity of the scoring system is supported, however, by several sets of evidence (see also the comments on normative responses on page 845). First, scores increased with age on both problems. Second, both scores correlated positively with verbal ability and with each other. Third, although sample size is clearly relevant to covariation judgments, it is uncommon for children, early adolescents, middle adolescents, and young adults to consider covariation information and sample size simultaneously. In general, adolescents place more emphasis on covariation data than on sample size (see, e.g., Jacobs & Narloch, 2001; Koslowski, Okagaki, Lorenz, & Umbach, 1989) and have much more relaxed standards for "adequate" samples than do trained statisticians (Klaczynski, 2001; Koslowski, 1996).

3. Knowledge calibration. Immediately after each selection task and each contingency task, predictions that solutions were correct were obtained on 5-point scales (1 = very certain, 5 = not at all certain). These certainty estimates index the metacognitive abilities to evaluate the extent to which one has particular cognitive skills, can generate logical products with those skills, and can synchronize this knowledge with performance.

For each deontic selection task, well-calibrated responses (scored +1) occurred when falsification solutions were coupled with "certain" or "very certain" ratings. Responses were overconfident (scored −1) when participants were "certain" or "very certain" of their responses but solved the problems incorrectly. Uncalibrated responses (scored 0) occurred when participants were uncertain that falsification responses were correct. Responses were also scored 0 when uncertainty ratings were coupled with incorrect (i.e., nonfalsification) solutions. Total deontic calibration scores could range from −2 to +2. An arbitrary calibration score was computed in the same way (−1, 0, +1); however, because only 6 of 142 participants were well calibrated on the arbitrary problem (in part because there were few falsification responses on this problem), this score is not further discussed. Although the nonnormative solution/uncertain rating combination could be considered well calibrated, the meaning of such responses is somewhat ambiguous. They may, in fact, have been given when participants understood task requirements and recognized that their responses were unlikely to fulfill those requirements. Alternatively, these responses could merely indicate that adolescents were reporting that they knew that their responses were guesses or that they simply did not understand the problem. Despite this ambiguity, these responses were considered superior to overconfident responses. The latter, but not the former, were related negatively to age and to ability.

A similar scoring system was used for the contingency detection problems. Responses were considered well calibrated (scored +1) when contingency relationships were identified accurately and respondents were "certain" or "very certain." Responses were poorly calibrated (−1) when respondents were confident that they had correctly identified the contingency but were, in fact, incorrect. Uncalibrated responses (0) were instances in which the contingency task was correctly solved but respondents indicated uncertainty, and when uncertainty ratings were coupled with incorrect solutions. Total contingency calibration scores could range from −2 to +2.

4. Statistical reasoning. These three problems were adapted from Stanovich and West (1998a). In each problem, a decision-making situation (e.g., to take a traditional lecture class or a computer-based class) was presented, along with two arguments. One argument contained large sample information (e.g., course evaluations from numerous students) that supported one decision. The second argument involved a small sample of personalized evidence (e.g., complaints of two honor-roll students) that favored the other decision. The argument supported by the large sample was counterbalanced across forms, as was the order in which the small- and large-sample decision preferences were presented.

After reading each vignette, participants decided which action to take on a 4-point scale (1 = definitely take the lecture course, 4 = definitely take the computer-based course). On a second 4-point scale, participants indicated whether the large-sample argument was more or less intelligent than the small-sample argument. On both the statistical decision-rating scale and the statistical intelligence-rating scale, higher ratings indicate more reliance on statistical evidence than on experiential evidence. Following Stanovich and West (1998a), ratings of 3 or 4 were deemed statistical and scored 1; ratings of 1 or 2 were scored 0. On both scales, total scores could range from 0 to 3.

Although the two scales were correlated, $r = .68$, there were numerous instances in which preference for the large-sample argument on one rating scale was coupled with preference for the small-sample argument on the other scale. In 84% of these cases, respondents acknowledged the large-sample argument as more intelligent, but based their decisions on the small-sample argument. These conflicts thus consisted of decisions based on a heuristic, such as "seeing is believing," and a contradictory intelligence rating. Total statistical conflict scores could range from 0 to 3.

5. Conjunction problems. On each of four problems—including the infamous "Linda" problem (see Gigerenzer, 1996; Kahneman & Tversky, 1996)—brief vignettes describing an individual's appearance, personal tastes, and other background information were presented. Each vignette was designed to cue beliefs that the individual belonged to each of two social categories (e.g., athlete and popular student). Six possible descriptors (in one of three randomly determined orders) of the individual were then presented. Two of these were distracter items (e.g., is a teacher's pet) and the third was the conjunction of the two distracters. The remaining three items depicted the two relevant social categories individually and as a conjunction. Items were rank ordered in terms of the likelihood that each accurately described the individual.

In any conjunction, p(AB: athlete and popular) cannot exceed p(A: athlete) *or* p(B: popular) because the individual categories (A and B) necessarily include all subcategories (e.g., athletes who are popular *must* be members of the superordinate category "athletes"). Therefore, normative conjunction responses involved rating each component of the conjunction as more likely than the conjunction.

Although solutions were scored as erroneous when conjunctions were rated as more likely than *either* of their components, an interesting NNR was the "strong conjunction fallacy," scored when conjunctions were rated as more probable than *both* of their components. Because such responses indicate inferences that the target matched stereotypical conceptions of both individual categories, they are assumed to be more indicative of the representativeness heuristic than responses in which conjunctions were rated as more likely than only one component. Normative conjunction scores and strong conjunction fallacy scores could range from 0 to 3.

6. Gambler's fallacy problems. In gambler's fallacy problems, predictions are made regarding brief sequences of events for which objective probabilities are supplied. In each of two problems, participants were given information about the likelihood of an event (e.g., 25% chance of winning a round of video poker) and information about a recent series of trials in which the event occurred more frequently than would have been expected from the objective probability (e.g., winning on six of eight trials).

A range of probabilities from 0 to 100 was presented. This range included the alternatives lower than the objective probability (e.g., 5%, 10%), equal to the objective probability (i.e., 25%), greater than the objective probability but less than the observed frequency (e.g., 35%, 50%), equal to the observed frequency (e.g., 75%), and greater than both the actual probability and the observed frequency (e.g., 90%). The task was to indicate the probability that the event would occur on the next trial. The NCR was selection of the alternative that matched the objective probability.

The compensation fallacy, an NNR, occurred when participants indicated that the likelihood of the event was lower than the objective probability—thus, the next trial would compensate for the prior sequence in which the event appeared to occur at a greater-than-expected frequency. Such responses may be generated by an averaging heuristic (e.g., "After an event sequence occurs at a greater-than-expected rate, subsequent events will occur at a lower-than-expected rate until the average is reestablished."). Normative gambler responses and compensation fallacy scores could range from 0 to 2.

7. Outcome bias. Outcome bias occurs when decisions are evaluated on the basis of their outcomes rather than on a priori probabilities of success or failure. Four hypothetical scenarios, two involving medical decisions and two involving economic decisions, were presented. For each decision domain, one scenario contained information that a decision had a relatively high probability of failure, but was nonetheless successful. The decision on the second problem in each domain involved a relatively low probability of failure, but which failed nonetheless. On Day 1, one problem from each domain was presented; the second problem from each domain was presented on Day 2. Outcomes and probabilities were counterbalanced across problems (e.g., the low probability with positive outcome on one form was the high probability with negative outcome on another form); no effects of form or presentation order (i.e., whether the high probability with negative outcome problem was presented as the first or second problem) were found.

The quality of each decision was rated on a 7-point scale (1 = very poor decision; 7 = very good decision). Outcome bias was determined by subtracting ratings on the high probability with negative outcome problems from ratings on the low probability with positive outcome problems. Summed across the two decision domains, bias scores could range from −12 to +12. Outcome bias was indicated by positive scores.

8. Hindsight bias. One part of the hindsight bias problem depicted two people who were currently romantically involved. The second part, presented separately, depicted two people who had been in a relationship that had dissolved. In each part, four pieces of information about the couple were provided, involving such matters as hobbies, shared interests, and former boyfriends or girlfriends. Two forms were created such that the parts involving the separated couple on one form involved the intact couple on the second form and vice versa.

On the post hoc part of the problem, respondents used a 5-point scale (1 = extremely unlikely; 5 = extremely likely) to indicate the likelihood that they would have been able to predict the dissolution of the couple. On the intact part of the problem, respondents indicated the likelihood that they could accurately predict whether the couple would remain intact in the future. Hindsight bias was calculated by subtracting ratings on the intact part from ratings on the post hoc part. Scores could range from −4 to +4; positive scores indicated hindsight bias.

Half of the participants received the intact part on Day 1 and the post hoc part of the problem on Day 2; this order was reversed for the remaining participants. No significant differences between presentation order or forms were found, which indicates that bias scores were not artifacts of the different contents of the two parts.

RESULTS

Results are presented in three sections. First, multivariate analyses of variance (MANOVAs), with age as the sole independent variable and scores on the various measures listed previously as the dependent variables, were conducted. These analyses were followed by univariate ANOVAs to determine the specific measures on which the early and middle adolescents differed. Next, correlations among measures were performed. The third section presents two factor analyses, each intended to explore latent variables that could explain the intermeasure correlations.

Age Differences in Normative and Nonnormative Responses

Mean scores for each age group on each measure are presented in Table 1. For ease of comparison, total deontic falsification (and deontic *pq*), normative contingency, statistical decision ratings, statistical intelligence ratings, statistical conflict, normative gambler's solutions, gambler compensation fallacy, normative conjunction, and strong conjunction fallacy scores are reported as proportions. Deontic calibration, contingency calibration, and outcome bias scores are reported as averages.

Initial analyses revealed no effects for gender or form, which were excluded from subsequent analyses. A MANOVA with deontic falsification, arbitrary *pq*, normative contingency, deontic calibration, contingency calibration, statistical decision ratings, statistical intelligence ratings, statistical conflict, normative conjunction, normative gambler's solutions, outcome bias, and hindsight bias scores as dependent variables indicated a significant effect for age, $F(12, 129) = 3.25$,

Table 1 Mean Normative and Nonnormative Responses, Separately for Early and Middle Adolescents

	Age Group	
Measure	Early Adolescent	Middle Adolescent
Deontic falsification[a]	.45 (.44)	.66 (.42)**
Deontic *pq*[b]	.14 (.29)	.05 (.15)*
Arbitrary falsification[a]	.03 (.17)	.12 (.33)*
Arbitrary analytic[a]	.12 (.33)	.25 (.38)*
Arbitrary *pq*[b]	.41 (.50)	.43 (.50)
Normative contingency[a]	.35 (.30)	.52 (.35)**
Deontic calibration[a]	.06 (.64)	.42 (.60)***
Contingency calibration[a]	−.15 (.54)	.13 (.51)***
Statistical decision ratings[a]	.18 (.39)	.42 (.50)**
Statistical intelligence ratings[a]	.23 (.42)	.45 (.50)**
Statistical conflict[b]	.28 (.27)	.17 (.19)**
Normative conjunction[a]	.29 (.30)	.26 (.29)
Strong conjunction fallacy[b]	.58 (.66)	.41 (.45)*
Normative gambler's solution[a]	.24 (.33)	.41 (.42)**
Gambler compensation fallacy[b]	.28 (.36)	.12 (.24)**
Outcome bias[b]	4.21 (3.02)	2.78 (3.07)**
Hindsight bias[b]	.32 (.91)	.59 (1.11)
Verbal ability	14.73 (5.13)	17.91 (5.61)***

Note: Scores are presented as averages; exceptions are verbal ability, the three scores derived from the AST, and hindsight bias, each of which was based on only one task. Standard deviations are in parentheses.
[a] Higher scores indicate more normative responding.
[b] Higher scores indicate more nonnormative responding.
* $p < .05$; ** $p < .01$; *** $p < .001$.

$p < .001$. Because deontic *pq*, arbitrary falsification, strong conjunction fallacy, and compensation fallacy scores were not independent of other scores, a separate MANOVA was conducted on these measures. Again, the effect for age was significant, $F(4, 137) = 5.51$, $p < .001$. Age effects remained significant in a pair of MANCOVAs with verbal ability as a covariate, $F(12, 128) = 3.22$, $p < .001$, and $F(4, 136) = 4.33$, $p = .006$, respectively. A final ANOVA was conducted on arbitrary analytic scores because these scores were not independent of arbitrary *pq* or arbitrary falsification scores.

Univariate analyses revealed age differences on 14 of the 17 measures (see Table 1). The hypothesized age-related increase in normative responding was supported: Middle adolescents had significantly higher scores than early adolescents on each of the variables that were considered normative, with the exception of normative conjunction responses. Age trends in the remaining responses were not as consistent, although the general pattern was toward age-related decreases in NNRs.

Although middle adolescents were more normatively competent than early adolescents, responding on most measures was poor relative to normative standards. For instance, only 42%, 29%, and 41% of the middle adolescents' responses to the statistical decision rating, conjunction, and gambler's problems, respectively, were normative. Most middle adolescents were poorly calibrated to their contingency judgments, 75% were biased by outcomes, and 46% displayed hindsight bias; only 25% met the relaxed criterion for normative responding on the AST.

Correlational Analyses

The next analyses examined relations among response categories. (Because they occurred infrequently, deontic *pq* and arbitrary falsification were excluded from these analyses. Statistical decision and intelligence ratings were similarly related to other responses and were, therefore, combined to create an average statistical ratings composite.) Because 12 categories were age related, age was partialled out of the correlations to decrease the likelihood that the observed patterns of covariation were artifacts of a general developmental variable (e.g., biological maturation, fluid intelligence) that age indexed.

In Table 2, the partial correlations among response categories are organized to highlight clusters of significant correlations. Two patterns are evident. First, there is considerable positive manifold among NCRs. Of the 28 correlations among deontic falsification, arbitrary analytic, normative contingency, deontic calibration, contingency calibration, averaged statistical ratings, normative gambler's responses, and normative conjunction responses, 22 (78.6%) were statistically significant.

The second pattern concerns variables labeled as nonnormative (i.e., statistical conflict, gambler compensation fallacy, strong conjunction fallacy, outcome bias, and hindsight bias). Again, considerable positive manifold is evident: Of 15 correlations, 11 (73.3%) were significant. Only the *pq* response on the AST was not linked to the other NNRs.

Less obvious are the relations among NCRs and NNRs. Excluding correlations between nonindependent variables (e.g., normative conjunction/conjunction fallacy, arbitrary analytic/arbitrary *pq*), only 10 of 45

Table 2 Partial Correlations among Response Categories, Controlling for Age

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Deontic falsification		.42****	.44****	.82****	.26****	.19*	.23***	.18*	−.12	−.16*	−.06	−.11	.01	.18*
2. Arbitrary analytic			.21**	.44****	.21**	.07	.12	.26****	−.11	−.12	.06	−.08	.01	−.35****
3. Normative contingency detection				.37****	.78****	.17*	.26****	.18*	−.07	−.13	−.13	−.24***	−.13	.05
4. Deontic calibration					.30****	.27****	.25***	.16*	−.12	−.08	−.03	−.07	−.03	.19*
5. Contingency calibration						.11	.16*	.10	−.07	−.02	−.12	−.26****	−.08	.12
6. Average statistical ratings							.19*	.10	−.06	−.03	−.03	.03	.15*	.02
7. Normative gambler								.01	−.03	.07	−.22**	−.25***	.00	.00
8. Normative conjunction									−.19*	−.39****	−.05	−.15*	−.14*	−.07
9. Statistical conflict										.23***	.20*	.28****	.23***	−.07
10. Strong conjunction fallacy											.26****	.18*	.25***	−.07
11. Gambler compensation fallacy												.52****	.42****	−.16*
12. Outcome bias													.31****	−.08
13. Hindsight bias														−.07
14. Arbitrary *pq*														
Verbal ability	.18	.12	.15	.14	.07	.11	.16	.20	.05	−.14	.08	.08	.04	.07

* $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$.

NCR–NNR correlations were significant. Further, even after controlling for age effects, five of the eight NCRs, but only one of the six NNRs, were related to verbal ability. These patterns are consistent with the two-factor hypotheses that (1) the analytic system assumes predominance when NCRs are generated; (2) the generation of NNRs is systematic, which suggests heuristic system predominance; and (3) although both systems were presumably active in all responses, when one system was predominant over the other, the products of that predominance were generally independent.

Factor Analytic Results

To further test the hypothesis that analytic processing was the predominant factor underlying NCRs and that heuristic processing was the predominant factor underlying NNRs, two principal components factor analyses were performed. The high deontic falsification/deontic calibration and normative contingency/contingency calibration correlations, however, inevitably led to separate factors that represented only these correlations. Consequently, these four variables were standardized and combined to create two new variables, metadeontic competence and metacontingency competence. Because separate analyses by age group indicated factors similar to those found in analyses that were pooled across ages, only the latter analyses are presented.

In the first analysis, six NCRs (i.e., metadeontic competence, metacontingency competence, arbitrary analytic, averaged statistical ratings, normative conjunction, and normative gambler's solution) and three NNRs (statistical conflict, outcome bias, and hindsight bias) were included. Three factors emerged. These three factors accounted for 55.7% of the variance shared among variables. Eigenvalues (and percent variance accounted for) for the first, second, and third factors were 2.46 (27.3%), 1.45 (16.1%), and 1.10 (12.3%), respectively. The internal consistency of the factors is indicated by the high loadings of the variables highlighted in Table 3. Four of the six NCRs loaded on the first factor (labeled Analytic-1), which supports the expectation that predominantly analytic reasoning enabled these responses. Further supporting this inference are the positive correlations of Analytic-1 scores with verbal ability and age.

The high loadings of the three NNRs on the second factor (labeled Heuristic) suggest that these responses had a shared origin in predominantly heuristic processing. This conclusion is bolstered by the nonsignificant correlations of Heuristic scores with verbal ability and age.

Table 3 Factor Analysis of Six Normative and Three Nonnormative Response Categories

Measure	Analytic-1	Heuristic	Analytic-2
Normative gambler's solution	<u>.72</u>	−.26	−.18
Averaged statistical ratings	<u>.63</u>	.19	.12
Metadeontic competence	<u>.60</u>	−.01	.53
Metacontingency competence	<u>.56</u>	−.28	.24
Outcome bias	−.27	<u>.75</u>	−.01
Hindsight bias	.30	<u>.72</u>	−.05
Statistical conflict	−.14	<u>.56</u>	−.20
Normative conjunction	−.15	−.21	<u>.74</u>
Arbitrary analytic	.31	−.03	<u>.71</u>
Criterion variable			
Verbal ability	.24*	.02	.27**
Age	.41**	−.09	.05

Note: The partial correlations between Analytic-1 and verbal ability and between Analytic-2 and verbal ability were .15, $p < .05$, and .26, $p < .001$, respectively, after controlling for age. With verbal ability controlled, the partial correlation between Analytic-1 and age was .37 ($p < .001$).
* $p < .01$; ** $p < .001$.

The two remaining NCRs (normative conjunction and arbitrary analytic) loaded on the third factor (labeled Analytic-2). Like the Analytic-1 factor, Analytic-2 was related to verbal ability; unlike Analytic-1, however, Analytic-2 was not related to age.

Because several NNRs (e.g., compensation fallacy scores) were not independent from several NCRs (e.g., normative gambler's solution scores), they could not be included in the first analysis. A second factor analysis was therefore conducted on three NCRs (i.e., metadeontic competence, metacontingency competence, and averaged statistical ratings) and six NNRs (i.e., statistical conflict, outcome bias, hindsight bias, arbitrary *pq*, strong conjunction fallacy, and gambler compensation fallacy). Three factors, which accounted for 55.8% of the variance among measures, were again indicated. Eigenvalues (and percent variance accounted for) for the first, second, and third factors were 2.45 (27.2%), 1.52 (16.9%), and 1.06 (11.8%), respectively.

As shown in Table 4, five NNRs loaded on the Heuristic factor, and the three NCRs loaded on the Analytic factor. The Heuristic factor was not associated with verbal ability, although it was (in contrast to the first factor analysis) related negatively to age. Analytic scores were linked to age and to verbal ability.

The isolated status of arbitrary *pq* responses can be explained in several ways. For example, scores were based on a single item and thus may have been unreliable. Similarly, the range of *pq* scores (0–1) was more restricted than the range for some of the other NNRs (e.g., outcome bias and hindsight bias). Alternatively,

Table 4 Factor Analysis of Three Normative and Six Nonnormative Response Categories

Measure	Heuristic	Analytic	*pq*
Gambler compensation fallacy	.77	−.07	−.06
Outcome bias	.74	−.16	.33
Hindsight bias	.66	.34	−.16
Statistical conflict	.53	−.21	−.01
Strong conjunction fallacy	.52	−.12	−.17
Metadeontic competence	−.13	.73	.27
Averaged statistical ratings	.06	.73	−.08
Metacontingency competence	−.29	.63	.01
Arbitrary *pq*	−.08	.09	.92
Criterion variable			
Verbal ability	.03	.33**	.10
Age	−.20*	.40**	−.05

Note: The partial correlation between Analytic scores and verbal ability was .27, $p < .001$, after controlling for age; partial correlations between Analytic scores and age and between Heuristic scores and age were .34, $p < .001$, and −.20, $p < .01$, respectively, after controlling for verbal ability.
* $p < .05$; ** $p < .001$.

of the NNRs assessed here, it could be argued that the *pq* response has been most convincingly established as heuristic (see Evans & Over, 1996) and may be a more purely heuristic response than the other NNRs.

DISCUSSION

The present research explored several questions not previously addressed by adolescent decision-making researchers. Key findings were as follows: (1) Normative responding increased and nonnormative responding decreased with age on most measures. (2) With age effects controlled, 35 of the 45 correlations between NCRs and NNRs were not significant. (3) In contrast, most normative responses were significantly interrelated; similarly, NNRs were generally related to one another. (4) In neither factor analysis did NCRs and NNRs load on the same factors. NCRs and the analytic factors were consistently related to age and, to a lessor extent, verbal ability. The heuristic factors were, at best, weakly related to age and ability.

Age, Processing Systems, and the Competence/Performance Issue

Several findings are unique to this study and shed new light on the development of adolescent decision-making. Outcome biases appeared to decline from early to middle adolescence but nonetheless remained ubiquitous. Hindsight biases and normative conjunction responses, in contrast, were not age related. The strong conjunction fallacy—which indicates considerable reliance on the representativeness heuristic—was, however, less common in middle adolescents. Middle adolescents were better able than early adolescents to accurately gauge the results of their reasoning, a developmental advance with a clear metacognitive component. Statistical reasoning improved with age, although most responses suggested use of a "seeing is believing" heuristic. Internal contradictions in statistical reasoning, which Sloman (1996) argues are strong indicators that two processing systems are operating simultaneously, indicated that adolescents, like adults, often decide against their better judgment (see Denes-Raj & Epstein, 1994).

Do age-related increases in normative responding and decreases in nonnormative responding indicate that analytic processing becomes predominant during adolescence? The analytic competencies for solving the problems in this study are typically acquired before middle adolescence (Inhelder & Piaget, 1958; Koslowski, 1996; Moshman, 1998). For example, probabilistic reasoning abilities—central to responding normatively to contingency, statistical reasoning, conjunction, gambler's, and outcome bias problems—are well developed by early adolescence (Kreitler & Kreitler, 1986; Piaget & Inhelder, 1951/1975; Reyna & Brainerd, 1994). Only on the deontic selection tasks and the contingency detection problems, however, did the "average" middle adolescent respond normatively. Table 1 shows that most adolescents do not demonstrate a level of performance commensurate with their abilities. If normative responses are used as an index, the data provide little evidence for the predominance of analytic processing. In experimental situations with well-defined tasks (relative to more commonplace tasks), middle adolescents are more likely to rely on analytic processing than early adolescents, but this is not their primary means of decision-making. Yet caution should be exercised before equating normative responses with analytical processing. Indeed, if some NCRs (e.g., DST falsification solutions) are heuristic products, then the data probably underestimate the extent to which adolescents typically engage in analytic processing.

The question is not, "Why do so few adolescents have the ability to solve these problems?", because the competence *is* available. The more important question is, "Can the gap between competence and performance be explained by traditional developmental theories?" Historically, treatments of competence/performance discrepancies have relied on evidence for familiarity effects (e.g., Ward & Overton, 1990), random errors (e.g., Stein, 1996), confusing item presentation, and so forth. Such (often ad hoc) explanations fail primarily because they do not explain the systematicity of and covariations among nonnormative responses.

Stanovich (1999) offers an explanation of between-ability level differences in responding that may also be applicable to age differences. In his view, ability differences are likely to arise under two conditions. In the first case, tasks are relatively barren of superficial content, clearly entail precise responding, and leave little room for misinterpretation (e.g., the inductive problem, "2, 5, 8, 11, ??"). Such tasks may be relatively "pure" measures of analytic competence (indeed, fluid intelligence tests presume such purity). If tasks meet these requirements, they would seem well suited to determining the developmental timetables of analytic competencies (although counterarguments are abundant; see Overton, 1990). As task difficulty increases, so too should between-age and between-ability differences—at least until tasks require operations beyond the average apex of development.

The more interesting case for decision-making researchers arises when tasks pit analytic processing against heuristic processing (Epstein, Lipson, Holstein, & Huh, 1992). Because illustrating this conflict for each task would be unduly cumbersome and extensive analyses of several tasks have been reviewed elsewhere (Stanovich, 1999), only a brief examination of outcome bias problems is presented here.

In a typical problem, a decision with a low (or high) probability of success is presented. If only that information were provided, most adolescents would probably opt against the decision. Analytic versus heuristic conflict is created, however, because the low probability information must be weighed against knowledge that the outcome was successful. The probability "pulls" for analytic processing and the outcome "pulls" for heuristic processing. The majority of adolescents committed the "contrary-to-fact" fallacy (e.g., "Monday morning quarterbacking") and evaluated decisions on an a posteriori basis.

To avoid this bias, adolescents must (1) understand basic concepts of probability (which most do), (2) appreciate the difficulties inherent in post hoc theorizing, and (3) decontextualize the logic of the problem (e.g., "Is success sufficiently likely to justify the decision?") from irrelevant content (the success or failure of the decision). The strong attraction of heuristic responding may lie in adolescents' experiences with decisions. In most situations, probabilities of success and actual success are congruent. Because outcomes are easier to process than probabilities, and because reliance on either usually leads to the same evaluation, adolescents focus on outcomes to judge decision quality. Middle adolescents may be less prone than early adolescents to this bias, not only because they are better able to inhibit memory interference (e.g., automated rules activated by heuristic processing, such as "successful outcomes = good decisions"), but also because they are more predisposed toward decontextualization.

The smaller competence/performance gaps among middle adolescents on other tasks may be accounted for by similar interplay between inhibiting memory interference (e.g., from normally useful rules that are susceptible to overgeneralization) and extracting decontextualized representations. Recall that most older adolescents displayed outcome bias and, on different tasks, other nonnormative tendencies. Critical to the further development of this account, therefore, are investigations of the phenomenological experiences associated with different tasks. Presumably, even if most adolescents experienced the hypothesized conflict, they were more strongly attracted to contextualized representations and to heuristic processing (e.g., because of speed and cognitive economy) than to decontextualizing tasks from content and to analytic processing.

The data do not entirely resolve the question of whether deontic falsification solutions are analytically or heuristically produced. The correlation between these responses and ability supports the view that this response is analytic, as does the strong correlation between falsification responses and certainty estimates. Specifically, if the responses were produced without awareness, then this correlation should have been weak and, possibly, negative. Despite this evidence for analytic predominance, the modest size of the falsification–ability correlation leaves open the possibility that some falsification responses were generated analytically, whereas others were generated heuristically. Further, as discussed subsequently, heuristically generated products may be available for conscious reflection; adolescents could thus be well calibrated to a solution without knowing precisely how they generated that solution.

Other issues were not entirely resolved by this research. For instance, despite increasing acceptance of two-process theories in social and cognitive psychology, theoretical accounts of the two systems are underspecified (Moshman, 2000). For instance, although the tendency to equate analytic processing with normative responding may be acceptable in exploratory investigations, many automatic inferences—for example, when preschool children make probabilistic and deductive inferences (e.g., Hawkins, Pea, Glick, & Scribner, 1984; Huber, & Huber, 1987)—are entirely compatible with normative standards. Similarly, nonnormative responses cannot be entirely equated with heuristic processing because such responses (like the "*p* only" response) may result largely from failures in analytic processing.

One position on this particular dilemma is that the heuristic processing/analytic processing distinction confounds conscious or unconscious inferential processes with heuristic or analytic inferential processes, and that this distinction places undue emphasis on the role of consciousness in normative responding (Chen & Chaiken, 1999; Moshman, 2000). This concern is well founded, but the proposed solution (i.e., four processes: conscious–analytic, unconscious–analytic, conscious–heuristic, unconscious–heuristic) is problematic. For example, the distinction between unconscious–analytic processing and unconscious–heuristic processing is at odds with the view of many two-process theorists that "quick, easy, and cognitively cheap" inferences are adaptive, and thus normative, under many conditions. Because some of these inferences are in accord with logical standards and others are not is insufficient reason to justify that they are generated by separate systems. At present, there is no evidence for the postulate that one unconscious processing system leads to automatic inferences that are consistent with traditional norms and that a separate unconscious processing system leads to automatic responses that are inconsistent with these norms.

A similar approach, advocated here, is to retain the current distinction between heuristic processing and analytic processing while simultaneously paying greater attention to the distinction between analytic and heuristic processes and the products of these processes. Heuristic processing and analytic processing both generate normative and nonnormative products. Further, some (but probably not all) products of the heuristic processing system may, at least momentarily, be available for conscious inspection. These products may be logical inferences or heuristic intuitions, such as representativeness. Once generated, they are sometimes applied immediately because of time constraints or because of their "feel-right" qualities. Before invoking them, however, reflective, metacognitively predisposed adolescents may evaluate these products and weigh them against analytically generated options.

This argument, then, is that the process–product relationship can be captured by a 2 × 2 classification scheme that distinguishes processes (predominant processing system: analytic or heuristic) from product (consistent or inconsistent with traditional norms). A rough overview of this scheme is presented in Table 5. It is important to note that there are some qualitative differences between some of the examples of analytic and heuristic products listed in the four cells. Specifically, many nonnormative heuristic products are holistic and imprecise, whereas nonnormative analytic products are precise and motivated by accuracy concerns. Although attention to the partial independence of process from product should help alleviate concerns that unconscious processing leads to maladaptive responding and that conscious processing leads to adaptive responding, further refinement of this ap-

Table 5 Proposed Processing System × Product Classification Scheme

	Processing System	
Product	Analytic	Heuristic
Normative	Logical deductions (e.g., falsification responses to the AST and DST) Mathematical computations Probabilistic inductions Decisions based on cost–benefit analyses or scientific reasoning Conversational inferences	Automatic logical inferences (e.g., possibly DST falsification) Automatic and gist-based mathematical solutions Decisions based on holistic situation estimates Conversational inferences Inferences based on facial cues
Nonnormative	Logical fallacies "*p* only" responses to the AST Computational/logical inferences on misrepresented problems Errors in mathematical computations Inaccurate cost–benefit analyses	Certain logical fallacies Biases (e.g., perceptual, belief based, self-serving, outcome, hindsight) Global heuristics (e.g., "seeing is believing," representativeness) Stereotype-based inferences

Note: Importantly, normative and nonnormative are used to refer to traditional standards and violations of those standards. Each of the products in the four cells may be normative under certain conditions and nonnormative under other conditions (for examples, see Evans & Over, 1996; Klaczynski, 2001; Stanovich, 1999). DST = deontic selection task.

proach is in order. For instance, current approaches make black-and-white distinctions between conscious and unconscious processing and ignore the possibilities that consciousness exists on a continuum (D. Moshman, personal communication, April 14, 2000), and that a degree of minimal awareness may sometimes be involved in heuristic processing.

The Development of Two Reasoning Systems

The results indicate that responses traditionally deemed normative were related to age and to ability and loaded on similar factors, which implies that the same latent variable can account for them. The findings support the expectation that this variable is an analytic processing system, which operates to decontextualize structure from surface-level content and to apply internalized rules of logic and decision-making. Analytic processing presumably entailed more cognitive effort and more thorough contemplation of task details than heuristic processing. The consequence of this more intense deliberation was better understanding of logical requirements and extraction of relatively content-free representations to which higher order reasoning skills were applied.

The exception to the systematic covariation among NCRs was the normative conjunction response. Insight into this unexpected finding comes from separate analyses of the early and middle adolescents. The NCR correlation matrix indicated that only 12 of the 28 correlations were significant for the early adolescents. Although the smaller sample can partially account for fewer significant correlations in these analyses than in Table 2, normative conjunction responses were not related to any other NCR, largest $r = .19$. In the middle-adolescent data, 20 of 28 correlations were significant; normative conjunction responses were significantly related to five of the other seven NCRs and were related to verbal ability only for the middle adolescents. Thus, the abilities responsible for NCRs, including those necessary for normative conjunction responses, appear to increasingly consolidate during adolescence.

This pattern of relations is consistent with that reported by Stanovich and West (1998a, 1998b, 1998c) in their investigations of college students. Those researchers also presented heuristics and biases tasks and found that NCRs (e.g., syllogistic and statistical reasoning problems) were positively correlated to each other and to general ability. Again, this positive manifold implies a well consolidated reasoning system (an analytic structure d'ensemble; see Flavell & Wohlwill, 1969; Moshman, 1977; Piaget, 1972) among middle and late adolescents.

In addition to replicating and extending the findings of Stanovich and West (1998a), this research also examined relations among nonnormative biases and fallacies. Because many of these had not previously been studied, the positive manifold observed among NNRs is particularly significant, as is the observation that these responses were related neither to ability nor to NCRs. The speculations that heuristic processing was predominant in producing these responses and that this processing system is partially independent from analytic reasoning were thereby supported. The additional finding that five of the six NNRs loaded on the same factor bolsters the conclusion that the underlying processing system was noncomputational in nature.

Separate analyses of the two age groups indicated that 7 of the 10 correlations in the NNR matrix (excluding arbitrary *pq* responses) were significant for both the early and the middle adolescents. Strategies produced by heuristic processing thus appear to be at least loosely consolidated by early adolescence. Unlike analytically produced responses, these may not consolidate further with age.

The results are relevant to recent discussions regarding the nature of rational thought. One school of thought, embodied in Stein's (1996) argument, is that apparently irrational thinking arises from random "slips" in reasoning that is otherwise rational. This account implies that average correlations should approach zero. Stanovich and West (1998a, 1998b, 1998c) provide initial evidence falsifying this line of argumentation among college students. The present research extends their findings to additional tasks and to adolescents. Together these investigations should dispel myths that nonnormative responding is random and arises from nonsystematic, mysterious, and otherwise inexplicable sources.

A second approach to rationality assumes that humans are maximally rational within limitations inherent to information processing, such as working memory size (Gigerenzer, 1993; Simon, 1993). If working memory capacity increases with age (Case, 1998), then the obvious developmental prediction is that irrational thinking should be less frequent in adolescence than in preadolescence and childhood. To some extent, the data support this prediction. Nonetheless, normative responding was modal on few tasks and did not increase with age on other tasks. Recent research with children, early and middle adolescents, and young adults shows that although the competence to respond normatively increases with age, these progressions are often not accompanied by parallel decreases in numerous nonnormative responses (e.g., Davidson, 1995; Jacobs & Potenza, 1991; Klaczynski & Narasimham, 1998b).

If the "bounded rationality" hypothesis is correct, it applies only to a subset of responses. Specifically, NNRs—which bounded rationality theorists sometimes argue are "satisficing" strategies when capacities are overwhelmed—were unrelated to verbal ability and, in one factor analysis, to age. If ability and age index computational capacity, significant negative relations should have been found. Further, if normative responding is limited primarily by capacity, correlations between NCRs on one task and NNRs on other tasks should have been consistently negative: Adolescents who had sufficient capacity to respond normatively to statistical reasoning and deontic selection problems should not have been overwhelmed by tasks that do not involve more complex operations (e.g., hindsight bias and outcome bias). Coupled with findings of memory and reasoning independence, the data indicate serious shortcomings in bounded rationality models.

Psychologists and philosophers interested in rationality have typically judged responses dichotomously, either as normative–rational or as nonnormative–irrational. Reyna et al. (in press) propose an alternative, process-oriented theory of rationality. In their view, nonnormative responses are judged at various grades of rationality. At the lowest level, NNRs occur because the requisite competence is lacking. Reasoning that operates on inappropriate gist representations is placed at a higher level. In the present vernacular, this level is akin to relying on contextualized, rather than decontextualized, representations. Although definitive conclusions await, most competence/performance gaps observed here appear to have arisen at this level. In the Reyna et al. view, however, still more rationality is seen when reasoners correctly represent tasks, but retrieve inappropriate principles or misapply appropriate principles—as may have been the case with "*p* only" responses to the AST.

An obvious advantage of a process-oriented approach is the notion that different mechanisms may underlie the same nonnormative response. On contingency detection problems, for instance, the NNRs of early adolescents may occur because they operate on the wrong representation (e.g., by neglecting denominators). Middle adolescents may give identical responses, but do so because they miscalculate ratios. This conceptualization of rationality invites researchers to move away from black-and-white classifications of performance. Rather than dubbing heuristic tendencies as inherently irrational, perhaps it is better to judge them as irrational only to certain degrees and only under specific conditions. The variability of "real-world" conditions, and the cognitive effort required, make maintaining a consistent level of higher order reasoning difficult, impractical, and, in some cases, counterproductive. Social conditions and personal goals fluctuate frequently such that analytic processing will be more fruitful in some contexts, and heuristic reasoning more suitable to other contexts. The emphasis on "appropriate" principles in this account is not identical with views that there is a response that should always be considered normative on a given task. Consistent with the present view, the "correctness" of a response or principle is context dependent; however, neither approach espouses an "anything-goes" relativism (see Reyna et al., in press).

The present research represents a step toward introducing two-process theories into cognitive developmental psychology. Additional research, with more diverse JDM tasks and with manipulations that induce analytic or heuristic predominance, is needed to determine the applicability of this approach to younger children and to achieve a better understanding of the relevant processes. Why are some tasks associated with such strong "heuristic attractions" that they overwhelm individual differences in age and ability? Do cultures differ in the extent to which they rely on one system or the other? When should heuristic responses be considered normative? How is cognitive maturity best defined? Questions such as these must be addressed before a comprehensive two-process theory of cognitive development can be constructed.

ACKNOWLEDGMENTS

The author would like to express appreciation to the students, teachers, and principals of Fairview Elementary School and Owen High School for their assistance in conducting this research, and to Sue Nations for her assistance in coordinating this project. Marcie Bean, Terry Sloan, and Danica Ciamareti deserve thanks for their assistance in data collection. The author is particularly indebted to Keith Stanovich and Valerie Reyna for raising many of the issues that framed this work. Portions of this work were presented at the Annual Conference of the Southeastern Psychological Association, Savannah, Georgia, 1999, and at the Annual Conference of the Society for Research on Adolescence, Chicago, 2000.

ADDRESS AND AFFILIATION

Corresponding author: Paul A. Klaczynski, Department of Psychology, Pennsylvania State University, Moore Building, University Park, PA 16802; e-mail: pak21@psu.edu.

APPENDIX

SELECTION TASK RULES

Deontic, familiar: If a person is drinking beer, then that person must be at least 21 years old.

Deontic, unfamiliar: If a person enters the country of XasDu, then that person must have at least two children.

Arbitrary: If the letter "A" is on one side of a card, then the number "7" must be on the other side of the card.

CONTINGENCY DETECTION

A doctor has been working on a cure for a mysterious disease. Many people are getting sick from the disease, so the doctor has been working very hard. Finally, he created a drug that he thinks will cure people of the disease. Before he can begin to sell it, he has to test the drug. He selected 14 people for his test and compared how sick these people were after getting the drug to 7 people who did not get the drug. Here's what he found.

	Cured	Still Sick
Got the drug	8	6
Did not get the drug	4	3

LAW OF LARGE NUMBERS

Ken and Toni are teachers who are arguing over whether students enjoy the new computer-based teaching method that is used in some math classes.

Ken's argument is, "Each of the 3 years that we've had the computer-based learning class, about 60 students have taken it. At the end of each year, they have written essays on why they liked or didn't like the class. Over 85% of the students say that they have liked it. That's more than 130 out of 150 students who liked the computer class!"

Toni's argument is, "I don't think you're right. Stephanie and John (the two best students in the school, both are high-honors students) have come to me and complained about how much they hate the computer-based learning class and how much more they like regular math classes. They say that a computer just can't replace a good teacher, who is a real person."

CONJUNCTION

Timothy is very good-looking, strong, and does not smoke. He likes hanging around with his male friends, watching sports on TV, and driving his Ford Mustang convertible. He's very concerned with how he looks and with being in good shape. He is a high school senior now and is trying to get a college scholarship.

Based on this information, which of the following statements is most likely to be true? Rank each statement in terms of how likely it is to be true, using a number from 1 to 6. The most likely statement should get a 1. The least likely statement should get a 6. [Note: The actual statements were presented in one of two randomly determined orders. Numbers of normative responses and strong conjunction fallacy responses did not differ across orders.]

____Timothy is popular
____Timothy is an athlete
____Timothy is an athlete and is popular
____Timothy has a girlfriend
____Timothy is a teachers' pet
____Timothy is a teacher's pet and has a girlfriend

GAMBLER'S FALLACY

When playing video poker, the average person beats the computer one in every four tries (25% of the time). Julie, however, has just beaten the computer six out of eight times (75% of the time). What are her chances of winning the next time she plays? (Circle one of the choices below.)

OUTCOME BIAS

High Failure Likelihood with Favorable Outcome

A businessman owned a company that was not making very much money. The man was, of course, very upset: If he did not make more money, he would be forced to shut the company down in the next 4 years. He learned that he could save his company if he became partners with another company. By joining together, he could make enough money to keep the company going at least 8 more years. However, there was a 10% chance that both companies would go bankrupt and lose all their money if he became partners with the other company.

The man decided to go ahead and become partners. The partnership worked; now the company will last at least 8 more years. Was the person's decision to become partners with the other company a good decision?

Low Failure Likelihood with Unfavorable Outcome

A man has been running a small market for 10 years, but the market has never made much money. He has learned that because of a new shopping mall, he will be forced to close his market within 3 years, but he could move his store into the mall. If he moved into the mall, he could keep his market open for at least 9 more years. However, there was a 4% chance that his store would completely fail if he moved it into the mall.

The man decided to go ahead and open a new store in the mall. It failed and he had to close his store. Was the man's decision to move his store into the mall a good decision?

HINDSIGHT BIAS

Intact Problem

Paul and Laura have been married for 1 year. Here are some facts about them.

- Both Paul and Laura like to hike and to travel.

- Paul enjoys watching sports, but Laura enjoys watching soap operas.
- Paul spends Friday nights at a bar with his friends and every Tuesday night he goes bowling with "the guys."
- Laura works as a secretary for the man she used to date.

Knowing these facts, how likely do you think it is that you could predict whether Paul and Laura will get divorced?

Post-hoc Problem

Randy and Amy just broke up. They had been a couple for 10 months. Here is some information about them before they broke up.

- Randy liked to drink beer with his friends.
- Amy liked to shop for clothes, but Randy hated shopping.
- Both Amy and Randy liked to work out and play racquetball.
- One of Randy's best friends was his old girlfriend.

Knowing these facts, how likely do you think it is that you would have been able to predict that Randy and Amy would break up?

REFERENCES

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than animals? *Psychological Review, 125*, 591–600.

Bargh, J. A., & Chartland, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462–479.

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88*, 463–502.

Brainerd, C. J., & Gordon, L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology, 30*, 163–177.

Brainerd, C. J., & Kingma, J. (1984). Do children have to remember to reason? *Developmental Review, 4*, 311–377.

Brainerd, C. J., & Kingma, J. (1985). On the independence of short-term memory and working memory in cognitive development. *Cognitive Psychology, 17*, 210–247.

Brainerd, C. J., & Reyna, V. F. (1992). The independence effect: What do the data show? What do the theories claim? *Developmental Review, 12*, 164–186.

Byrnes, J. P. (1998). *The nature and development of decision making: A self-regulation model.* Hillsdale, NJ: Erlbaum.

Case, R. (1998). The development of conceptual structures. In D. Kuhn & R. Siegler (Eds.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 745–800). New York: Wiley.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York: Guilford.

Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy in children's decision making. *Merrill-Palmer Quarterly, 41*, 328–346.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology, 66*, 819–829.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49*, 709–724.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62*, 328–339.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive-experiential self-theory. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462–482). New York: Guilford.

Evans, J. B. St. T. (1989). *Bias in human reasoning: Causes and consequences.* London: Routledge.

Evans, J. B. St. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87*, 223–240.

Evans, J. B. St. T., & Over, D. E. (1996). *Reasoning and rationality.* Hove, U.K.: Psychology Press.

Flavell, J. H., & Wohlwill, J. F. (1969). Formal and functional aspects of cognitive development. In J. H Flavell & J. F. Wohlwill (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget.* New York: Oxford University Press.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. Manktelow & D. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 284–313). London: Routledge.

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

Hawkins, J., Pea, R. D., Glick, J., & Scribner, S. (1984). "Merds that laugh don't like mushrooms": Evidence for deductive reasoning by preschoolers. *Developmental Psychology, 20*, 584–594.

Hayes, B. K., & Hennessy, R. (1996). The nature and development of nonverbal implicit memory. *Journal of Experimental Child Psychology, 63*, 22–43.

Huber, B. L., & Huber, O. (1987). Development of the concept of comparative subjective probability. *Journal of Experimental Child Psychology, 44*, 304–316.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

Jacobs, J. E., & Narloch, R. (2001). Children's use of sample size and variability to make social inferences. *Journal of Applied Developmental Psychology, 22*.

Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*, 166–178.

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

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

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

Klaczynski, P. A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development, 71*, 1347–1366.

Klaczynski, P. A. (2001). Framing effects on adolescent task representations, analytic and heuristic processing, and decision making: Implications for the normative-descriptive gap. *Journal of Applied Developmental Psychology,* 22.

Klaczynski, P. A., & Narasimham, G. (1998a). The development of self-serving reasoning biases: Ego-protective versus cognitive explanations. *Developmental Psychology, 34*, 175–187.

Klaczynski, P. A., & Narasimham, G. (1998b). Problem representations as mediators of adolescent deductive reasoning. *Developmental Psychology, 34*, 865–881.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.

Koslowski, B., Okagaki, L., Lorenz, C., & Umbach, D. (1989). When covariation isn't enough: The role of causal mechanism, sampling method, and sample size in causal reasoning. *Child Development, 57*, 1316–1327.

Kreitler, S., & Kreitler, H. (1986). Development of probability thinking in children 5 to 12 years old. *Cognitive Development, 1*, 365–390.

Lie, E., & Newcombe, N. S. (1999). Elementary school children's explicit and implicit memory for faces of preschool classmates. *Developmental Psychology, 35*, 102–112.

Manktelow, K. I., Sutherland, E., & Over, D. E. (1995). Probabilistic factors in deontic reasoning. *Thinking and Reasoning, 1*, 201–219.

Markovits, H., & Dumas, C. (1999). Developmental patterns of understanding social and physical transitivity. *Journal of Experimental Child Psychology, 73*, 95–114.

Mitchell, P., Robinson, E. J., Isaacs, J. E., & Nye, R. M. (1996). Contamination in reasoning about false belief: An instance of belief bias in adults but not children. *Cognition, 59*, 1–21.

Moshman, D. (1977). Consolidation and stage formation in the emergence of formal operations. *Developmental Psychology, 13*, 95–100.

Moshman, D. (1998). Cognitive development beyond childhood. In D. Kuhn & R. Siegler (Eds.), W. Damon (Series Ed.) *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 947–978). New York: Wiley.

Moshman, D. (2000). Diversity in reasoning and rationality: Metacognitive and developmental considerations. *Brain and Behavioral Sciences, 23*, 689–690.

Muller, U., Sokol, B., & Overton, W. F. (1998). Reframing a constructivist model of the development of mental representation: The role of higher-order operations. *Developmental Review, 18*, 155–201.

Newcombe, N., & Fox, N. A. (1994). Infantile amnesia: Through a glass darkly. *Child Development, 65*, 31–40.

Overton, W. F. (1990). Competence and procedures: Constraints on the development of logical reasoning. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 1–32). Hillsdale, NJ: Erlbaum.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15*, 1–12.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* New York: Norton. (Original work published 1951)

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239–272). New York: Wiley.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75.

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science, 5*, 275–279.

Reyna, V. F., Lloyd, F. J., & Brainerd, C. J. (in press). Memory, development, and rationality: An integrative theory of judgment and decision-making. In S. Schneider & J. Shanteau (Eds.), *Emerging perspectives on decision research.* Cambridge, MA: Cambridge University Press.

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

Schneider, W., & Bjorklund, D. F. (1998). Memory. In D. Kuhn & R. Siegler (Eds.), W. Damon (Series Ed.) *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 467–521). New York: Wiley.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Simon, H. A. (1993). Decision making: Rational, nonrational, and irrational. *Educational Administration Quarterly, 29*, 392–411.

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

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

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). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161–188.

Stanovich, K. E., & West, R. F. (1998b). Who uses base rates and P(D/-H)? An analysis of individual differences. *Memory and Cognition, 26*, 161–179.

Stanovich, K. E., & West, R. F. (1998c). Cognitive ability and variation in selection task performance. *Thinking and Reasoning, 4*, 193–230.

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

Thurstone, T. G. (1962). *Primary mental abilities for grades 9–12.* Chicago: Science Research Associates.

Ward, S. L., & Overton, W. F. (1990). Semantic familiarity, relevance, and the development of deductive reasoning. *Developmental Psychology, 26*, 488–493.

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology.* Harmondworth, U.K.: Penguin.

Wasserman, E. A., Dorner, W. W., & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 509–521.