Psychological Bulletin
1989, Vol. 105, No. 3, 331–351

0033-2909/89/$00.75

Approaches to Studying Formal and Everyday Reasoning

Kathleen M. Galotti
Carleton College

Reviews the current literature on different types of reasoning, showing little integration across different studies and very little explicit discussion regarding the nature of reasoning. Moreover, most recent theoretical work on deductive and inductive reasoning does not make any explicit connection to everyday reasoning. Major programmatic approaches to the study of reasoning are classified into three types: the componential approach, the rules/heuristics approach, and the mental models/search approach. Strengths and shortcomings of each are reviewed. It is concluded that each approach can potentially play an important role in describing one or more aspects of everyday reasoning. In terms of direct extendability to the description, study, and improvement of everyday reasoning, the mental models/search approach is argued to be the most useful, especially when it incorporates a generalized version of the rules/heuristics approach.

The study of thinking and reasoning has expanded and intensified in the past two decades. The emergence of the fields of cognitive science and artificial intelligence have reawakened interest in the problem of describing, in a useful and testable way, what makes for "good," "critical," "reflective," and "productive" thinking (Dewey, 1933; Wertheimer, 1945). Programs and studies to assess and improve critical thinking are currently underway in a number of educational contexts (Baron, 1985; Baron & Sternberg, 1987; Neimark, 1987; Nickerson, Perkins, & Smith, 1985; Perkins, 1985a, 1986a; Raths, Wasserman, Jonas, & Rothstein, 1986; Vye, Delclos, Burns, & Bransford, 1988). However, despite many impressive findings, unresolved fundamental issues and problems threaten to impede substantial future progress. One of the most important, I argue, is how well current models of specific reasoning tasks describe and explain everyday reasoning performance.

Many of the problems stem from a lack of agreement over definitions of terms. Boundaries for usage of the term *reasoning* are unclear, leaving it very difficult for a reader to know if two investigators, each purporting to study a kind of reasoning, are really studying a common entity. It is hard to meaningfully interpret patterns of performance across different laboratory reasoning tasks. Moreover, investigators tend to develop models that apply only to a narrow range of tasks. It is generally assumed that the specific task (or range of tasks) underlies or bears some other important relationship to other reasoning tasks and to thinking and reasoning in everyday life. However, the relationship between the laboratory task and everyday reasoning in general is never explicated in any detail, leaving open the question of whether the task being studied is really of significant interest.

In this review, I examine three major programmatic approaches to the study of reasoning that appear to have generality that extends beyond one specific task: the componential approach, the rules/heuristics approach, and the mental models/search approach. My purpose is to assess the strengths and weaknesses of each approach in accounting for performance in a variety of contexts and on a variety of tasks, both laboratory and everyday. Before turning to these three approaches, I first take up matters of definition, attempting to draw some boundaries around the term *reasoning.*

Reasoning Versus Thinking and Decision Making

Of the three terms, *thinking* is meant to be the most general and, on probably any account, to include reasoning and decision making, although the lines among the three are blurred (see Johnson-Laird [1988] and other chapters in Sternberg & Smith [1988] for a variety of points of view on the interrelations of these terms). *Thinking* has been defined as "going beyond the information given" or "doing anything at all with evidence" (Bruner, 1957), filling gaps left by incomplete information (Bartlett, 1958), or "searching through a problem space" (Newell, 1981; Newell & Simon, 1972). When used in contrast to reasoning, thinking is meant to cover other kinds of intellectual activity such as problem solving, decision making, or brainstorming. *Problem solving* typically refers to people's behavior in the laboratory, when confronted with a situation that requires one or more insights to solve. *Decision making* refers to an assessment of, and choice from among, alternatives in terms of their probability of occurrence and their expected value. *Brainstorming* usually refers to the generation of ideas used to solve either a laboratory or an everyday problem.

Some theorists distinguish between good thinking and "biased," or "unproductive," thinking. It is when distinctions between productive or critical and nonproductive or noncritical thinking are drawn that the boundaries between thinking and reasoning become most blurred. Wertheimer (1945), in distin-

Support for this work was received from a Faculty Development Grant from Carleton College. I am grateful to Lloyd Komatsu for providing valuable advice, extensive criticism, and technical assistance, and to Jonathan Baron, Deborah Appleman, Ulric Neisser, and Robert Sternberg for comments and suggestions on earlier drafts.

Correspondence concerning this article should be addressed to Kathleen M. Galotti, Department of Psychology, Carleton College, One North College Street, Northfield, Minnesota 55057-4025.

guishing productive thinking from "blind induction," equated the former with grasping the essential structure of the problem. Dewey (1933), describing reflective thinking, viewed it as a process used when no easily applied formula or rule exists to reach a goal, resolve a state of doubt, or decide on a course of action. Ennis (1987) defined critical thinking as "reasonable reflective thinking that is focussed on deciding what to believe or do" (p. 10). For both Ennis and Dewey, good (i.e., reflective or critical) thinking involves dispositions (e.g., to be open-minded) as well as skills. I argue here that critical thinking, in its usual usage, is really an instance of everyday reasoning.

Some theorists make little or no distinction between thinking and reasoning. Boole (1854/1951, cited in Dellarosa, 1988), describing a system of propositional reasoning, entitled his seminal work, "An investigation into the laws of thought." Others who have studied so-called weak (i.e., domain general as opposed to domain specific) strategies of thinking and problem solving would also probably be disinclined to draw strong distinctions between thinking and reasoning.

There are various proposals for what such domain-general strategies are (see Lesgold, 1988, for a more complete review). One, a means–ends analysis, conceives of the thinker as analyzing a problem into one or more goal states, the current state, and a set of operators that bring about transformations in states. Thinking consists of a set of selections of operators, together with the corresponding transformations in states, which lead from the current state to the goal state or states. Good thinking is that which constructs an efficient path from the current to the goal state or states. The generate-and-test strategy, a second general strategy of thinking, consists of two phases: the listing of candidate solutions to a problem (generation), followed by evaluation of these candidates (testing). Good thinking here is construed as a combination of effective and focused generation and of thorough and rigorous testing.

Other theorists consider reasoning to be a particular type of thinking. However, reasoning, too, has been described in a rather broad way:

> There is of course no clear boundary surrounding this topic [reasoning]. It is obvious, for example, that when an individual draws a conclusion from premises according to traditional Aristotelian laws of logic, he is engaged in reasoning. It is also feasible to assert that an individual solving a crossword puzzle, planning to buy a new house, or determining the best route from one town to another, is also engaged in reasoning. (Wason & Johnson-Laird, 1972, p. 1)

Those psychologists who consider reasoning as something distinct from other types of thinking typically study performance on a number of well-defined self-contained tasks, such as verbal analogies, categorical syllogisms, conditional syllogisms, linear syllogisms, or series completion problems (Anderson, 1985). Sometimes, the term *reasoning* is restricted to intellectual activity with problems whose solutions are governed by a system of logic, such as propositional or predicate calculus (Braine, 1978; Braine, Reiser, & Rumain, 1984; Osherson, 1975; Rips, 1983, 1984).

Investigators of decision making emphasize it as a process of assessing probabilities, predicting values of outcomes, and using various decision rules (Kahneman & Tversky, 1982a; see Kahneman, Slovic, & Tversky, 1982, for a review). At the same time, decision making clearly draws upon the same skills used in reasoning and in other types of thinking. Inferences are involved in gathering and evaluating the data on which to make a choice or commitment. Therefore, I argue that instances of decision making regarding important life choices also ought to be treated as instances of everyday reasoning.

At first blush, it might seem most prudent from a theoretical perspective to adopt the clearest and most narrowly circumscribed definitions. One could restrict *reasoning* to mean "thinking according to the theorems of a logical system"; *decision making* to mean "weighting and combining probabilistic information in such a way as to rank alternatives"; and *thinking* to cover both of these, as well as other tasks in which given information is used or combined to lead to new information. If generally adhered to, this move might lead to clearer communication among cognitive psychologists and researchers in allied fields.

However, such a move would raise serious problems. In drawing these distinctions, one loses the ability to explain the intuition that all three terms refer to very similar kinds of intellectual activities. Further, if reasoning were to be defined narrowly as solving problems by the use of one or more systems of logic (deformed, incomplete, or otherwise), then the relationship between reasoning and everyday functioning, in which rules of logic may not always be applicable, is obscured. In turn, a critical question would arise: Why study reasoning? If solving a categorical syllogism has little to do with thinking outside of the laboratory, then what is the point of constructing a complex theory of it?

One might try to justify the study of syllogistic reasoning on the sole grounds that measures of syllogistic reasoning correlate with scores on psychometrically based intelligence tests (see Snow, Kyllonen, & Marshalek, 1984). However, note that intelligence test scores are of theoretical interest because they are supposed to predict everyday intellectual functioning (i.e., thinking outside the laboratory). Therefore, even on this account, the justification for study of syllogistic reasoning must eventually be grounded in its (syllogistic reasoning's) relationship to everyday reasoning, and researchers studying syllogistic reasoning eventually incur the burden of specifying and explaining that relationship.

To extend research on reasoning to examples that are typical of ordinary life, one needs to develop relatively broad definitions of reasoning. At the same time, one needs to guard against being unclear about the phenomenon under study in the name of being inclusive. It is all too easy to leave large chunks or models or theories undefined or partially specified and then, with hand waving, to claim great generality, leaving for others the Herculean task of establishing how the model constructed for one task applies to another.

Even within the existing circumscribed formal reasoning literature, generalizability of models to tasks other than the one or ones originally studied is a problem. Consider, for example, the following abbreviated descriptions of models of reasoning performance:

1. *On the process of reasoning with two- and three-term series problems (i.e., linear syllogisms):*

> Its identifiable stages are: (a) comprehension of the propositions; (b) comprehension of the question; (c) search for information

> asked in the question; (d) construction of an answer. (Clark, 1969, p. 392)

2. *On the process of reasoning with verbal analogies:*

> The subject begins analogy solution with attribute identification: He or she encodes the first analogy term, . . . and then the second. . . . In encoding each analogy term, the subject identifies the term, retrieves from long-term memory the attributes that may be relevant for analogy solution, retrieves from long-term memory a value corresponding to each attribute, and stores the results as an attribute-value list in working memory. . . . Next the subject infers the relation between all values of corresponding attributes in the first two analogy terms. The relation is stored as a list of attributes with corresponding values in working memory. Attribute comparison in inference is exhaustive with respect to the encoded attributes. . . . The subject then encodes the third analogy term . . . enabling him or her to map the relation between the first and third analogy terms. Mapping, like inference, is exhaustive with respect to the attributes stored in working memory. (Sternberg, 1977a, p. 355–357, Model 1)

3. *On the process of reasoning with categorical syllogisms:*

> The model consists of four processing stages: (a) an encoding stage, in which the individual premises are given a first reading and in which the processes necessary to establish a data base operate; (b) a composite stage, in which logical operations work to produce a single predicate, representing the information in the premises; (c) a conclusion-encoding stage, similar to the premise encoding stage; and (d) a comparison stage, in which the composite information is compared with the contents of the conclusion. This last stage also includes a decision substage, in which the reasoner selects his response. (Revlis, 1975, p. 98)

At first glance, the models would seem to share many similarities: In each, reasoning is composed of a sequence of steps. All of the models describe the first step as encoding (assuming that comprehension in the first example is equivalent to encoding in the other two), and all involve the manipulation of information during processing. But upon closer inspection, it is evident that the similarities are a function of the fact that the descriptions use identical or similar terms to describe unspecified key component processes, such as encoding, comparing, or inferring. Because it is not clear in any of these models exactly what inferring, for example, really is, it is not possible to assess whether like processes are indeed being proposed. And without this assessment, it is not possible to move any closer toward a clear sense of what reasoning really is.

This complaint is hardly new. Goldman and Pellegrino (1984), for example, described the problem in the following way:

> Most process models for cognitive tasks describe performance in terms of the same basic or elementary information processes such as encoding, response, inference, or comparison. Psychologists currently use the same limited set of process labels to describe performance on a wide range of simple and complex tasks and this particularly reflects an assumption that there is a core set of elementary information processes. Unfortunately, little has been done to support or document their existence and generality. (p. 190)

Lest the three models just described be unfairly singled out, it is worth noting that this complaint can be applied to virtually any theory of reasoning that does not describe in detail all of the component processes, as well as to any computer model of reasoning that has any kind of wishfully labeled "black box" process (McDermott, 1981).

Here, then, is the major ambition: to develop a theory of reasoning that is specific enough to yield testable predictions but abstract enough to be applicable to a variety of tasks. The theory, like any good theory, must set clear boundaries around what it is and is not meant to explain. For example, it should be clear whether solving verbal analogies is to count as an instance of reasoning and why. It must spell out the components it uses to explain performance in some detail. At the same time, it must avoid explaining performance in terms of components that apply only to one or a limited number of tasks.

As a first pass at defining reasoning, in a more comprehensive way, I propose the following: mental activity that consists of transforming given information (called the set of premises) in order to reach conclusions. This activity must be focused on at least one goal (but may be focused on more than one). The activity must not be inconsistent with systems of logic when all of the premises are fully specified, although there may not always be an applicable system of logic to govern specific instances of reasoning. The activity may or may not be self-contained; that is, people may implicitly or explicitly add to, subtract from, or otherwise modify any or all of the premises supplied. When original premises are modified, the final conclusion must be consistent with the modified premises. The activity may, but need not, involve the breaking of mental set. The conclusions may, but need not, be startling or nonobvious at the outset of the activity. The conclusion may, but need not, be deductively valid.

Given this definition, laboratory tasks—including propositional and syllogistic reasoning, solving analogies and series completion problems, assessing probabilities, and rank ordering expected values of outcomes—are all proper instances of reasoning. The first two are more prototypical, by virtue of their self-contained nature, the clear existence of a governing system of logic, and the explicit presentation of the premises. More everyday tasks are potential instances of reasoning to the extent that they involve more than a momentary intuitive response. Thus, tasks that involve, for example, identification of assumptions (i.e., premises) and explicit consideration of the steps used in moving from assumptions to conclusions count as examples of reasoning. Clear examples of such tasks include evaluating arguments and constructing and testing (formal or informal) hypotheses. For example, in evaluating an argument, one moves from the premises supplied in the argument to an assessment of its overall strength or compellingness. There is in this example a clear goal: constructing an overall assessment. Depending on the specifics of the argument, rules of inductive strength or deductive validity may be invoked. If the reasoner modifies the original set of premises by, for example, thinking of related premises or constructing counterexamples, then the overall assessment of the quality of the argument should (if the person is reasoning) reflect those modifications.

What sorts of mental activities are excluded by such a definition? Primarily, this rules out any thinking that consists of momentary, intuitive, responding. Similar to Anderson's (1985, pp. 199–200) definition of problem solving, this definition excludes any one-step mental process as an instance of reasoning. Sudden flashes of insight, if indeed they are instantaneous (Perkins, 1981, offered compelling counterarguments), are thus excluded. So-called gut reactions or evaluations, when left unanalyzed (or when claimed to be unanalyzable), fail to count as reasoning. Responses not involving the transformation of information (e.g., simple memory retrievals) also do not count, nor does daydreaming or other forms of free association, because these activities lack a goal or focus.

How does this definition compare with previous ones? At first glance, the present definition appears more cumbersome and less focused. Wason and Johnson-Laird (1972), in a classic work, described reasoning as the extent to which a person, given a set of assertions, can "appreciate all that follows from them by virtue of logic alone, and remain unseduced by plausible, but fallacious conclusions" (p. 2). Evans (1982), circumscribing his focus to deductive reasoning, defined it in the following way: "A deductive reasoning task involves making an inference from information which is *given*. If the task requires access to memory of things which are not presented, then it is simply not a reasoning task. . . . For [this reason], the interesting area of pragmatic inference is not included" (p. 5). Sternberg (1986b) provided a process-oriented definition of reasoning, very much in keeping with his componential analysis of laboratory reasoning tasks (reviewed later): "Reasoning . . . [is] the controlled and mediated application of three processes—selective encoding, selective comparison, and selective combination—to inferential rules" (p. 281). Holyoak and Nisbett (1988), describing inductive reasoning, saw it as involving "the use of both rules about events in particular domains . . . and higher-order, more abstract rules termed *inferential rules*" (pp. 50–51).

The present definition, then, is one that is more inclusive than those given by Evans (1982) or Wason and Johnson-Laird (1972), who focused strictly on deduction, or that given by Holyoak and Nisbett (1988), who concentrated on induction. This definition is similar in scope to Sternberg's (1986b) but more explicitly includes nonlaboratory kinds of reasoning. It also makes fewer claims than does either Sternberg's or Holyoak and Nisbett's definitions regarding the number and types of cognitive processes or other mental structures that define reasoning.

Formal Versus Everyday Reasoning

The laboratory–everyday distinction (borrowed and adapted slightly from Perkins, 1986b), is similar to one made by Bartlett (1958), who categorized thinking into two different types: closed system and adventurous. *Closed-system thinking* referred to thinking about numerical and logical problems, in which all terms were definable in advance. *Adventurous thinking* included scientific, popular, and socialized thinking, in which the domain of thought was less bounded. Bartlett argued that the two types of thinking were similar in structure and in purpose (both filled gaps left by other information) but that the tactics of the two types of thinking differed. For example, Bartlett saw closed-system thinking as consisting of a sequence of steps but saw the process of everyday thinking (a kind of adventurous thinking) as more akin to pattern recognition.

Bartlett's closed-system thinking maps directly onto what I call *formal reasoning,* the solving of problems for which all premises or given information is specified in advance. Formal reasoning would include all problems of logic (propositional, predicate, modal, and other types, although psychologists have traditionally restricted attention to the first two), geometric analogies, series completion problems, and other problems similar in scope. The critical feature here is that all the information that the subject is to consider is explicitly set forth in the problem; the subject need not search memory or consult outside resources to look for other relevant premises. Many of the studies within cognitive psychology that involve thinking and formal reasoning fall within this category.

Informal or everyday reasoning, in contrast, covers the intellectual activities that compose the thinking done in our everyday lives: planning, making commitments, evaluating arguments, discovering and choosing options. In this type of reasoning, premises are usually not completely supplied with the problem. The reasoner is forced to search for relevant information, and indeed often faces, as a subproblem, the task of determining just what information is relevant.

In some ways, the formal–everyday distinction maps onto the distinction between well-defined and ill-defined problems (Glass, Holyoak, & Santa, 1979). Well-defined problems have the given information, the legal moves or operations, and the goal state or states fully specified; ill-defined problems leave one or more of these relatively unspecified. Most everyday reasoning problems are ill defined to some degree, and formal reasoning problems, by definition, are well defined.

Few researchers have studied everyday reasoning, probably because of the lack of an established, appropriate methodology. The pragmatic issues raised by the study of this type of reasoning seem overwhelming, especially in contrast to the relatively few practical difficulties of studying, for example, categorical syllogisms. In the latter case, the experimenter supplies the premises and even controls the order and duration of presentation. The experimenter also can indirectly manipulate, by changing the content of the problem while keeping the structure constant, the reasoners' level of interest and emotional reaction to the problem. With everyday reasoning problems (e.g., deciding whether to buy a house or evaluating the coherence of a political party's platform), very few of these matters are under the experimenter's control. Reasoners will use their world knowledge, perhaps to different extents. It may be impossible to create structurally identical versions of the same problem (for example, What decision is structurally identical to the decision of whether to have children?). In addition, many formal reasoning problems, because of their self-contained nature, often take little time to complete. Moreover, formal reasoning problems often (if not always) have answers that are clearly correct; everyday problems often do not, because the set of relevant premises cannot always be circumscribed in real-life problems. Hence, performance on formal reasoning problems is almost always easier to measure and to compare to prescriptive or normative models (Baron, 1985).

If everyday reasoning is to be conceived of as a different entity from formal reasoning, on what is the distinction to hinge? Several related proposals have been offered. Perkins (1986b) for example, drew the distinction in the following manner:

> The contrasting structures of informal and formal argument bring with them contrasting challenges. The developer of an informal argument must take care to consider multiple lines of argument on

Table 1
Differences Between Formal and Everyday Reasoning Tasks

Formal	Everyday
All premises are supplied.	Some premises are implicit, and some are not supplied at all.
Problems are self-contained.	Problems are not self-contained.
There is typically one correct answer.	There are typically several possible answers that vary in quality.
Established methods of inference that apply to the problem often exist.	There rarely exist established procedures for solving the problem.
It is typically unambiguous when the problem is solved.	It is often unclear whether the current "best" solution is good enough.
The content of the problem is often of limited, academic interest.	The content of the problem typically has potential personal relevance.
Problems are solved for their own sake.	Problems are often solved as a means of achieving other goals.

> both sides of the case, explore thoroughly his or her knowledge base for whatever seemingly unrelated knowledge might apply, doubt the premises and examine with more than a critical eye inferential steps. Facing these difficulties, the developer of an informal argument is relieved of another—ferreting out a line of argument that begins with axioms quite remote from the proposition in question and only gradually converges on it. Such long-chain arguments are relevant in the formal domain and building them is a distinct art in itself. (p. 4)

Scribner (1986) distinguished between *practical thought,* described as "thinking embedded in the larger purposive activities of daily life . . . that function to achieve the goals of those activities," and *theoretical thinking,* defined as "isolated mental tasks undertaken for themselves" (p. 15). Resnick (1987a), reviewing a number of studies, echoed many of these ideas, seeing "nonschool learning" as more functional and goal-directed, contextualized, socially shared, and situation specific than school learning (see Sternberg & Caruso, 1985, and Wertsch, Minick, & Arns, 1984, for similar arguments).

Table 1 summarizes the major differences between formal and everyday reasoning, as defined here, at least insofar as agreement can be reached over terms (I suspect, for example, that everyday reasoning is a subset of practical thinking or nonschool learning, but that debate appears premature). Note that the entries in Table 1 focus on descriptions of tasks. This leaves somewhat open the question of whether the cognitive processes called upon by the two types of task are similar or not.

Many (although not all) who study formal reasoning would probably claim to be doing so because they expect it to predict ability at everyday reasoning. Viewed from this perspective, formal reasoning investigations would be seen as exploring, under more controlled circumstances, the skills and abilities that occur in everyday life. However, not everyone would expect this relationship between formal and everyday reasoning. In fact, there exist at least three distinct views on the relationship between them.

The first is the view that formal reasoning is a part of everyday reasoning: "Syllogisms also appear implicitly in normal English prose. Of course, in natural contexts, the premises and conclusions aren't labeled, but the underlying structure is the same" (Halpern, 1984, p. 63). A related position is adopted by Wason and Johnson-Laird (1972), who distinguished between "pure" and "practical" reasoning. I assume that they used *pure* in a way similar to my use of *formal.* They argued that pure reasoning, being concerned only with truth-functional issues, is a subset of practical reasoning, which must concern itself with both truth-functional and causal issues.

On this account, formal reasoning should be easier than, or should overstate, one's ability to reason about real-world problems for two reasons: (a) In the reasoning problems that one faces daily, one has to recall or generate one's own premises. Presumably, once one has generated all the premises one will use, the actual reasoning processes are the same as those used in a formal reasoning task. This in turn implies that everyday reasoning is structurally equivalent to formal reasoning, but adds extra steps: generating and evaluating premises. The difficulty of these extra steps is nontrivial, and so everyday reasoning is distinctly more difficult. (b) In everyday reasoning, one may have certain emotional attachments to certain premises or to certain possibilities or conclusions. Overcoming the effects of these attachments (again, a problem that normally does not arise in formal reasoning tasks) requires extra effort. In summary, there is usually more information to work with in a real-life problem than there is with a categorical syllogism or verbal analogy, and there may often be more information to ignore. Therefore, drill and practice on formal reasoning problems ought to enhance reasoning more generally. Resnick (1987b) cited examples of exactly this view, although she herself generally rejected it. However, some existing educational programs, such as Lipman's *Philosophy for Children* (see Lipman, 1987), seem to incorporate such a view.

A second view, in some ways the converse of the first, is that the two types of reasoning share similar processes but that formal reasoning requires more work and thus is harder. One could argue that in restricting attention to a small set of premises, all given in the problem, the reasoner must compartmentalize knowledge, ignoring personal information or memory or certain kinds of heuristics. Such ignoring and compartmentalizing might in fact make the task a harder one than the kind of tasks faced ordinarily in the course of a day. In addition, in formal reasoning tasks, the reasoner must often adopt an analytical approach, ignoring "invited inferences" (Geis & Zwicky, 1971); taking care not to rely on ordinary language comprehension processes (Braine, 1978) or conversational maxims (Grice, 1975); and on some problems, holding to a standard of strict logical necessity, in contrast to a standard of what is pragmatically likely (Henle, 1962, 1978; Johnson-Laird, 1982). Johnson-Laird (1982) argued that everyday reasoning involves implicit inferences that depend upon general knowledge and generally go beyond the strictly necessary conclusion; in contrast, formal reasoning "depend[s] on the *further* ability to search for alternative models that violate putative conclusions" (Johnson-Laird, 1982, p. 1, emphasis added). Notice that in this view, it is formal reasoning that requires extra work.

A third and very different conception of the relationship between formal reasoning and informal reasoning is that the two call upon very different processes and share few similarities. Perkins (1986a, 1986b) adopted this view, as one sees in the earlier quotation, arguing that although the two types of reasoning share some structural features, they also differ structurally in fundamental ways. He presented three specific contrasts between formal and informal reasoning (Perkins, 1986a, pp. 196–197): (a) Formal arguments have a "long chain" structure, that is, many individual steps, each leading to the next. Informal arguments have a "fork" structure, that is, several short lines of argument, each one with some degree of uncertainty, but which all converge on the conclusion. (b) Formal arguments argue only one side of a case, because their logical validity ensures that a contrary argument on another side will be invalid; informal arguments typically must include arguments both pro and con, because both types of arguments usually can be made. (c) Formal arguments occur in a closed world, taking the premises for granted and not relying on any other information, whereas informal arguments can, in principle, draw data from any source. On this account, the motivation for studying formal reasoning has a more limited scope: This ability might be used to predict other, specific abilities, for example, mathematical or spatial reasoning ability, but should not be expected to index reasoning skills in general. However, this restriction in applicability ought to bring about, it seems, a corresponding restriction in current interest in formal reasoning tasks.

Major Findings in Research on Formal Reasoning

A review of the many findings that concern people's reasoning would be voluminous. Indeed, even if the focus is restricted to the four most well-studied reasoning tasks—linear series, propositional, syllogistic, and analogical reasoning—the review would still be beyond the scope of this article. For such a review, the reader is referred to Anderson (1985), Baron (1988), Evans (1982), or Matlin (1983). Notice that even such a broad review would still be incomplete, focusing only on a handful of reasoning tasks, all of them formal. My focus here instead is to characterize a short list of generalizations that can be made from these studies. These generalizations will be helpful when comparing different theoretical models of formal and everyday reasoning.

Findings With Schooled Adults in Western Cultures

The first generalization is that people have great difficulty reasoning with negative premises or with marked as opposed to unmarked adjectives. For example, categorical syllogisms containing premises such as "No A are B" or "Some A are not B" take longer, and often result in more errors, than do syllogisms that consist of premises that are positive (e.g., "Some A are B" or "All A are B"). Likewise, when reasoning with a linear series problem, premises, such as "Anne is better than Mike," that contain the unmarked term *better* are easier to reason with than are premises, such as "Mark is worse than Susan," that contain *worse,* a marked term (Matlin, 1983).

A second finding from studies of both propositional and syllogistic reasoning is that reasoners, in encoding premises, often inadvertently alter the premises' meaning. A premise of a categorical syllogism such as "All A are B" has been argued to be illicitly converted by some reasoners to also mean "All B are A" (Revlis, 1975). Some errors in drawing conclusions, Revlis (1975) believed, can be traced back to such misencodings of the original premises. In other words, the original premises given in the problem are not always the premises used in the actual reasoning. This problem seems to persist even when people are explicitly trained in interpreting single premises (Galotti, Baron, & Sabini, 1986).

Related to this problem is a third, one that can often lead to error in propositional or syllogistic reasoning. This problem is the intrusion of world or background knowledge, even when the problem calls for analytical reasoning from only the premises presented. Subjects untutored in formal logic especially have a difficult time reasoning without drawing on their personal knowledge bases. This problem is especially pronounced with concrete contents of problems, when the use of background knowledge is invited. This so-called content effect refers to the fact that people's performance with a given reasoning problem can be dramatically affected by what the problem is nominally about and can differ greatly on two different versions of the same problem. As an example, consider the following two syllogisms: (a) "Some A are B. Some B are C." (b) "Some men are dog owners. Some dog owners are women." Although the two problems are structurally identical, subjects are generally much more likely to come to the correct conclusion (viz., that no deductively valid conclusion exists) with the second problem (see Wason & Johnson-Laird, 1972, for other examples with propositional reasoning).

A fourth problem, again especially evident in propositional and syllogistic reasoning, is the failure to consider all possible interpretations of a premise. The premise "If A then B" allows the following three states of affairs to be true: A true and B true; A false and B true; A false and B false. The premise "All A are B" allows for the possibility of Bs that are not As (i.e., that B is a superset of A) in addition to the more common interpretation that A and B are sets that overlap completely. It is argued that some errors in reasoning come about because of a failure to consider the multiple possibilities allowed by the premises (Erickson, 1978; Johnson-Laird, 1983).

In addition to failing to consider all possible interpretations of premises, people often fail to consider more than one or at most a few ways of combining premise information. This means that in general, only a few possibilities allowed by the premises, and therefore only a few possible conclusions, are thought of. In searching for or generating evidence for a particular hypothesis or putative conclusion, people in general show what has been called "confirmation bias" (Mynatt, Doherty, & Tweney, 1977; Wason, 1977); that is, people tend to examine only evidence that supports their initial beliefs and to ignore information that contradicts it or that supports other possible beliefs. This failure, often termed a bias in thinking, is often also held to be a major shortcoming of reasoning, especially everyday reasoning (Baron, 1985, 1988; Nisbett & Ross, 1980; Perkins, 1985a).

Finally, and no doubt in part because of the preceding list of shortcomings, people are overconfident in the reliability and validity of their own judgment. People often cease working at a reasoning problem before the optimal amount of time (as would be judged by the individual upon reflection) has been invested (Baron, 1985); judge too highly in relation to the objective odds the probability of their conclusion being correct (Lichtenstein, Fischhoff, & Phillips, 1982); and exhibit a hindsight bias, believing in retrospect that events were more certain to occur than they actually were (Fischhoff, 1982).

Findings From Cross-Cultural Studies

Despite the many errors made, the studies in the preceding section also demonstrate people's competence at some formal reasoning problems. This competence, however, has often been found to be specific to people educated in the Western system. Scribner (1977) reported, for example, that adults from nonliterate backgrounds scored only slightly better than chance when asked to reason with syllogisms such as "All people who own houses pay house tax. Boima does not pay a house tax. Does he own a house?". Large discrepancies in performance between schooled and unschooled adults appear, not only in Scribner's work, but in almost all cross-cultural work investigating formal reasoning (see Cole & Scribner [1974] for a review of seminal studies and the Laboratory of Comparative Human Cognition [1983] for a more recent review).

What is it that unschooled adults do when confronted with formal reasoning tasks? Mary Henle's (1962) description of "failing to accept the logical task" is the most apt. These adults alter, omit, or add to the premises supplied, in order to make the problems more consonant with their own knowledge bases and beliefs about the nature of the task. Indeed, in a study in which adult students and nonstudents were asked merely to remember (as opposed to solve) syllogisms, students again out performed nonstudents, who often distorted the premises.

None of these results are to suggest that nonschooled adults are incapable of reasoning; indeed, from their own "translation" of the premises, they reason elegantly and logically. The point here is that the genre of formal reasoning—which calls for an analytic orientation to the text, which is about problems whose content may be far-removed from one's own interests, and which may contradict one's own background knowledge—is one that appears to require literacy (Olsen, 1986).

Reasoning, then, when indexed by performance on formal reasoning tasks, must be understood with respect to the goals and comprehension of the reasoner. The goals and comprehension, in turn, appear to be strongly influenced by literacy, other experience with schooling, and possibly, other cultural variables (see Buscaglia, 1987, for other examples). At least some of the variables that influence reasoning, then, have to do with the context in which the task occurs.

Findings From Developmental Studies

How and when do formal reasoning abilities develop? Again, a detailed review is beyond the scope of this article (see Braine & Rumain, 1983, and Falmagne, 1975, for reviews). However, a few general points and illustrations can be offered. First, under special and usually limited conditions, younger children show at least some deductive and inductive reasoning competence (Hawkins, Pea, Glick, & Scribner, 1984; Holyoak, Junn, & Billman, 1984; Shapiro & O'Brien, 1970). Second, skill at reasoning within literate cultures does indeed develop and appears to become fully functional only after early adolescence (Barratt, 1975; Galotti & Komatsu, 1989; Kuhn, 1977; Moshman & Franks, 1986; Overton, Ward, Novek, Black, & O'Brien, 1987; Roberge, 1970; Roberge & Craven, 1982; Staudenmayer & Bourne, 1977; Taplin, Staudenmayer, & Taddonio, 1974). Various accounts of what specifically accounts for developmental changes on specific reasoning tasks have been offered, and they can roughly be classified into four types (although the types are by no means mutually exclusive).

The first holds that what develops is a change in a child's or adolescent's definition of terms within a problem. This account is typically offered to account for developmental differences on deductive reasoning problems, such as conditional or categorical syllogisms. Taplin et al. (1974), for example, suggested that nine-year-olds equate if–then constructions with the biconditional syllogism ("if and only if") and that only older children (aged 13) and adults correctly treat it as a conditional syllogism. Neimark and Chapman (1975) made related claims regarding the interpretation of particular quantifiers (e.g., "some") in solving categorical syllogisms.

A second general proposal for what develops locates the source of change in an increasing understanding of, and appreciation for, task demands. Hawkins et al. (1984) found that even preschoolers could solve conditional syllogisms and could give theoretical justifications (as opposed to empirical ones, in which the child offers evidence from her or his own knowledge base) under constrained conditions, that is, when the content of the problems cued them. Thus, four- to five-year-olds given fantasy-content problems (e.g., "Bangas are purple animals. Purple animals always sneeze at people. Do Bangas sneeze at people?") before other types of problems (e.g., "Rabbits never bite. Cuddly is a rabbit. Does Cuddly bite?") showed a better level of performance, and gave more sophisticated justifications, than did same-aged children who received problems in a different order. The authors argued that when fantasy problems come first, preschool children are cued to avoid their proclivity to reason pragmatically, that is, to incorporate any or all of their world knowledge. Bucci (1978) anticipated this proposal, arguing that in children of all ages there is a " 'pragmatic processing' mode that becomes less obligatory with age. In pragmatic interpretations, meaning is determined by previously known factual relations between the things which words represent, rather than by grammatical relations between the words themselves" (p. 55).

A third general view is that what develops is a change in strategy toward the task and, in particular, a recognition that there may be more than one possible interpretation of, or solution to, the given problem (Acredolo & Horobin, 1987; Johnson-Laird, Oakhill, & Bull, 1986). Acredolo and Horobin (1987) gave 20 relational reasoning tasks (e.g., regarding transitive inferences about relative sizes of objects) to children in first, third, fifth, and sixth grades. Only the sixth graders showed "a ready facility for detecting the possibility of more than one correct solution to multiple solution problems" (Acredolo & Horobin, 1987, p.

13). Sternberg and Rifkin (1979) offered a related explanation, that is, that younger children are more likely to execute some cognitive operations in a self-terminating fashion, whereas older children are more exhaustive.

A fourth account posits capacity increases, particularly in entities such as working memory or attention span, as responsible for developmental increases in formal reasoning specifically and in thinking skills more generally (see Smith, Serta, & Gattuso, 1988, for a review). Finally, a fifth account, somewhat related to the one mentioned earlier, centers around the Piagetian idea of a mental reorganization during adolescence, one that moves the reasoner toward a more powerful logical system, known as formal operational thought. Hallmarks of formal operational thought are the following: the capacity for hypotheticodeductive thought, the use of propositional logic, the ability to separate form from content, the capacity for combinatorial reasoning (i.e., to lay out all possibilities in a systematic fashion), and the capability of seeing reality as only one possibility (see Gruber & Vonèche [1977] and Keating [1980] for detailed reviews on the nature of formal operational thought). For many reasons, both empirical and logical, the strict Piagetian idea of formal operations has not held up well (see Keating, 1980, for an extended discussion), although it appears clear that some sort of blossoming of formal reasoning ability does take place during the adolescent years.

Note that just about any combination of these explanations is tenable. Changes in understanding of terms within problems could be coupled with changes in strategies used in a formal reasoning task could result from a shift in underlying cognitive structures, or both. Adopting an analytical, rather than a pragmatic, approach to the task may accompany other shifts in attitude toward the task and may encourage reasoners to examine other, less obvious, possibilities. In any case, proficiency on formal reasoning tasks does improve with age, becoming more apparent on a wider range of tasks and being especially pronounced on problems that call for multiple interpretations.

It is clear from this discussion that many of the errors in reasoning occur selectively on reasoning tasks within the laboratory. However, it is also at least plausible that the sources of error on formal reasoning tasks also contribute to errors in reasoning in everyday life. Confirmation bias and overconfidence, for example, are easy to imagine in a real-life context (see Nisbett & Ross, 1980). On the other hand, incorrect or incomplete interpretations of premises, at least on the face of it, are harder to transport to models of everyday reasoning, primarily because few premises are "given" in everyday reasoning.

Major Findings in Research on Everyday Reasoning

The literature on everyday—or real-life, informal, or pragmatic—reasoning (the terms are often used synonomously, and I will treat them this way here), is small in extent. Studies on reasoning outside the laboratory are few and typically very recent. Therefore, the generalizations drawn in this section are necessarily tentative in nature.

I concentrate on four studies of nonlaboratory reasoning, selected to show a range of everyday reasoning abilities (but see Resnick, 1987a, and Sternberg & Wagner, 1986, for others). Ceci and Liker (1986) studied experts and nonexperts handicapping horse races. Lave, Murtaugh, and de la Rocha (1984) investigated peoples' arithmetic computations, both on school-like tests and in actual practice while grocery shopping. Wagner and Sternberg (1985, 1986) examined the simulated on-the-job reasoning of three groups of people: academic psychologists, business managers, and bank managers. And Perkins (1985a, 1986b) studied the reasoning of students (high school through graduate school) and adult nonstudents about a variety of aesthetic, social, and political propositions.

The typical results from these studies are that various ability measures of the different kinds of everyday reasoning correlate slightly, if at all, with traditional measures of intelligence (e.g., Wechster Adult Intelligence Scale scores or scores on other IQ-like tests). In some cases, the ability measures do correlate with other things. Expert handicappers in the Ceci and Liker (1986) study, for example, were more likely than nonexperts to rely on a multiplicative model of six variables when handicapping (nonexperts treated each variable independently). Wagner and Sternberg (1985, 1986) reported that professional success (see their articles for how this was assessed) was uncorrelated with scores on a standard verbal reasoning task, but it did correlate significantly with the amount of tacit knowledge, that is, the practical and goal-oriented intuitions about the factors (and their relative weightings) that contribute to successful functioning.

These findings hint that formal and everyday reasoning abilities may be orthogonal or at least may be influenced by very different kinds of factors. After all, formal reasoning tasks are typically included on the standard intelligence measures used in the studies just mentioned. However, it is worth noting, first, that this conclusion has not as yet been directly or definitively tested. For one thing, intelligence tests do not consist entirely of formal reasoning, and so that ability has not yet been isolated. Other potential problems (if they exist) may have to do with the nature of some of the measures themselves: IQ measures are typically derived from paper-and-pencil or verbal responses; everyday reasoning measures are often derived from extensive interviews or behavioral measures. Fredericksen, Carlson, and Ward (1984) argued that a verbal–behavioral difference in measures can be sufficient to produce different outcomes. Second, the studies that have been cited here have sometimes relied on small sample sizes and have used dependent measures of unknown psychometric properties (see Detterman & Spry [1988] and Ceci & Liker [1988] for a spirited exchange on these points). Thus, at the very least, one would want to see more replications on larger samples before drawing firmer conclusions.

As a very preliminary generalization, it appears that some of what predicts good everyday reasoning is the breadth and depth of the knowledge base (Wagner & Sternberg, 1985, 1986; see also arguments by Fredericksen, 1986). However, another important aspect seems to be process based: Good everyday reasoning includes a more thorough and less biased search for arguments and evidence (Perkins, Allen, & Hefner, 1983) or a more sophisticated integration of information (Ceci & Liker, 1986).

In the next section, I consider three programmatic approaches that exist at a level of generality to be potentially applicable to different kinds of reasoning, both formal and everyday.

These are the componential approach, the rules/heuristics approach, and the mental models/search approach. I first focus on the distinctiveness of each and then discuss the implications of each as a means of understanding both formal and everyday reasoning. Next, I consider more broadly the contrasts among the three models, discussing the particular strengths and weaknesses of each in accounting for the set of existing findings described earlier. The ultimate goal of this examination is to assess the extendability of the three approaches to everyday reasoning.

Three Programmatic Approaches to the Study of Reasoning

Componential Approach

The idea of analyzing a reasoning task into its component processes has been widely used by, and is most closely associated with, Robert Sternberg (1977a, 1977b, 1982a, 1982b, 1983, 1984, 1986a, 1986b), although others have also made important contributions within this area (see Pellegrino & Glaser, 1980, for a review). The general idea dates back at least to J. P. Guilford's (1967) structure of intellect model. The aim is to discover and specify the basic cognitive processes that are used by an individual in any particular task and then to account for individual differences and within-subject differences across tasks in terms of some aspect of those component processes (e.g., the probability of using a particular component at all, of using that component correctly or completely or both, and of acquiring the component). The approach assumes that the existence and use of components is fairly stable, although not unmalleable; thus the componential approach is related to, but distinct from, a capacities approach, which explains mental functioning in terms of fixed entities such as the size of working memory or speed of processing (Case, 1974; Jensen, 1982, 1984).

Sternberg's pioneering contribution to this approach came from his studies of reasoning with verbal analogies of the form "Washington:one::Lincoln: a. five b. twenty" (after Sternberg, 1986a). Sternberg described the following as necessary steps of the reasoning process: encoding of the terms; inferring the relationship between the A and the B term (in the above example, between *Washington* and *one*); mapping the relationship just found onto the new domain introduced by the C term of the analogy (here, *Lincoln*); and applying the relationship previously inferred in this domain to generate the best possible completion. In case neither of the provided alternatives match the answer generated, the reasoner additionally needs to justify the choice of the imperfect answer.

Each of the terms used to describe these steps is used to name a component (a performance component, to be more specific; see the next paragraph). Sternberg attempted to isolate each component and to study its parameters (e.g., duration, or probability of execution) by presenting some analogies at different levels of deformity, for example, by precuing some of the terms. As an example, on some trials, subjects would see *Washington: one::* before the start of the timed trial; results of such deformed trials (when compared to those of nondeformed ones) were used to estimate the amount of time it takes to encode two terms and to infer the relationship between them.

In more recent work (Sternberg, 1982b, 1984, 1986a), Sternberg distinguished among three types of components: performance components, metacomponents, and knowledge acquisition components. Performance components are those such as the ones described in the preceding paragraph, processes used in actually carrying out a cognitive task. Metacomponents are executive or higher order processes used in the planning and monitoring of a task, such as selecting lower order components, selecting a strategy for combining lower order components, or allocating attentional resources. Finally, knowledge acquisition components are used in learning new information. Examples include selective encoding (sifting relevant from irrelevant information), selective combination, and selective comparison (Sternberg, 1984, pp. 284–285). Every component has associated with it three important parameters: *duration* (i.e., how long it takes to execute), *difficulty* (i.e., probability of being executed erroneously), and *probability of execution* (i.e., the likelihood that it will be executed at all).

Using data from hundreds of trials per subject, Sternberg estimated these parameters for each performance component, in order to predict either response latency or error rate. In general, for any particular task (e.g., verbal analogies, geometric analogies, series completion problems), his models account for a truly large proportion of the total variance, usually in the .80s or .90s. Moreover, Sternberg argued that this approach can do more than explain performance on any individual task. He made two important claims: (a) This approach can help account for individual differences, and (b) it can explain the similarities of performance on distinct reasoning tasks (Sternberg, 1984, 1986b; Sternberg & Gardner, 1983).

The componential approach can explain individual differences in several ways. Subjects who perform less well could lack a component altogether, have the component but not use it, use it but in the wrong way, use it for an insufficient amount of time, use the wrong combination of components or the right combination of components but in the wrong order, or some combination of any of these. Sternberg cited findings that he believed instantiate some of these accounts, for example, that good analogical reasoners spend more time encoding than do poorer ones or that more intelligent subjects (as measured by psychometric intelligence testing) spend relatively more time on higher order planning (Sternberg, 1984).

The second proposed virtue of the componential approach is its ability to explain performance across a variety of cognitive tasks. Sternberg and Gardner (1983) found that performance on analogies, series completion, and classification problems was highly correlated. This suggests that a common model of processing can describe performance on all three of these inductive reasoning tasks. High correlations between tasks are presumed to arise from the fact that the two tasks call upon overlapping sets of performance components (and perhaps metacomponents). The componential approach can also potentially account for *g*, the entity postulated to explain the positive correlations of performance on apparently disparate intellectual tasks. Sternberg (1982b) explained these correlations by claiming that the metacomponents he proposed are "rather general across executions of a variety of tasks" (p. 436). Thus, even seemingly very different tasks call upon similar sets of metacomponents.

Some work on the effects of practice and training of component processes has been reported (Alexander, Haensley, Crimins-Jeanes, & White, 1986; Davidson & Sternberg, 1984; see Vye et al. [1988] for a more general review and Nickerson et al. [1985] for a review of related programs, such as Feuerstein's Instrumental Enrichment Program). Results generally suggest that moderate amounts of practice with specific components (Davidson & Sternberg [1984] singled out selective encoding, selective combination, and selective comparison; Feuerstein's program centers around slightly different processes) does increase performance on standard intelligence tests, including problems of both inductive and deductive reasoning, as well as on other transfer problems (e.g., mystery problems that require a subject to determine how a detective could have identified a criminal). Note, however, that so far as I know, no direct investigation of the relationship between training of component processes and everyday reasoning about personally relevant issues has been undertaken.

With regard to the list of common findings in formal reasoning presented in the section, "Major Findings in Research on Formal Reasoning," the componential approach appears best able to explain the findings that have to do with problems in encoding. These would include the finding that people have difficulty reasoning with premises that contain negative information or marked terms and the fact that people untutored in logic often misunderstand the meaning of premises, even after careful explanations are made. Presumably, a componential approach would argue that negative information and marked terms require more resources to process, thus lowering the overall probability of successful reasoning. Parameters on the encoding component include ones indicating the probability of execution and the probability of error in execution, both of which theoretically handle these encoding problems well.

The componential approach does less well at explaining content effects and the intrusion of background knowledge into reasoning, at least without some extra assumptions and parameters. It could explain the failure to consider all possible premise interpretations, or all possible combinations of premise interpretations, and could explain overconfidence phenomena with reference to both problems in encoding, or problems in the metacomponents, which direct the assembly of lower order components. But in order to do so, this approach would need to be more specific about how encoding components work and about what metacomponents are involved and how they function.

Despite the breadth and ambitiousness of Sternberg's program, a number of problems or ambiguities have become apparent. I list four of them here, in increasing order of seriousness. First, Sternberg's claims for generality of the approach as yet constitute a blank check. He claimed, for example, that "linear syllogisms require at least some of the same performance components as do related kinds of problems, such as categorical syllogisms. Both require encoding of premise information, decoding of negations, combination of information from pairs of premises, and response" (Sternberg, 1982a, p. 269). However, the specificity of this proposed analogy is unclear. Is the process of encoding the same in both tasks? The same name is given to the initial process of both tasks, but there is no compelling reason to believe that those initial processes are identical.

This point relates to the second major criticism, the black box nature of the components. Neisser (1983) argued that the performance components postulated by Sternberg that have any generality have no definition and those that are well defined lack generality: "Despite his claims of generality, [Sternberg] is content to model tasks one at a time, inventing components *ad hoc* as they are needed" (Neisser, 1983, p. 195). The inference component, named in such a way as to suggest its centrality to thinking and reasoning, is never "unpacked," other than to suggest that it involves a "discovery of one or more relations between objects or events" (Sternberg, 1986a, p. 91). Without further specification, the name may be a misnomer. Certainly, this definition of inference seems dissimilar from the inference process or processes postulated by Braine (1978) and Osherson (1975) when discussing propositional reasoning, that is, the use of modus ponens and other rules of logic. In any case, this definition of inference does not set clear boundaries on what is and what is not to count as an example. Sternberg and Gardner (1983) acknowledged that "a more nearly complete account of the commonalities in inductive reasoning (and in g) would require further unpacking of these component constructs" (p. 111).

Perhaps the most striking shortcoming of the componential approach centers upon its lack of applicability to real-life thinking or reasoning, despite Sternberg's exhortations to study intelligence in meaningful contexts (Sternberg, 1984). In one of his earliest papers, Sternberg began by positing the centrality of analogical reasoning to everyday cognition: "We reason analogically whenever we make a decision about something new in our experience by drawing a parallel to something old in our experience. When we buy a new pet hamster because we liked our old one or when we listen to a friend's advice because it was correct once before, we are reasoning analogically" (Sternberg, 1977b, p. 99). In another paper (Sternberg, 1977a), he postulated that the following performance components are necessary ingredients in reasoning with verbal or geometric analogies: encoding, inferencing, mapping, applying, and justifying. But it is never clear how such components would apply to his initial examples of hamster buying or advice taking. When making a decision about the hamster, for example, there are no prepackaged terms to encode. One might supply one's own premises, presumably recalling aspects of an old pet hamster, but such operations must more properly be described as retrieval, not encoding. Nor is it at all clear how components such as mapping or inferencing would really apply, partly because of the lack of definition of mapping or inferencing, but at least equally as much because the relationships between laboratory reasoning and real-world reasoning are never explored in any depth.

Pellegrino and Lyon (1979) made the same point, arguing that the generality of the componential method depends on two criteria: additivity and separability. *Additivity* refers to a property of a task such that it can be deformed in such a way as to estimate parameters for various component processes. The resulting task variations "must be *unaltered* versions of those in the main task, so that their characteristics do not change when added together" (Pellegrino & Lyon, 1979, p. 183, emphasis theirs). In addition, the task variations that are constructed must be such that they minimize the confounding of different parameters. This requirement is known as the *separability* criterion. At this point, it is not at all clear that everyday reasoning tasks can be found that meet the additivity and the separability requirements.

Neisser (1983) argued a related point even more strongly: Sternberg's experimental procedures (precuing and what Neisser called the "excessive" number of trials per subject: 2,880 in one case) force subjects to adopt a task-specific strategy that gives the impression that they solve analogies componentially. (It should be noted that other reviewers, such as Pellegrino & Lyon [1979], have faulted Sternberg on occasion for not gathering enough data to estimate stable parameter values.) Neisser, then, would maintain that Sternberg's findings bear no relationship to the reasoning that occurs in everyday experience. At the very least, the relationship between the laboratory performance and everyday reasoning has yet to be discussed at a useful level of detail.

To his credit, Sternberg maintained that his approach is a working one that is subject to revision, and he is committed to refining and extending the model to cover new instances of reasoning. Meanwhile, there are open issues regarding how many of these problems are solvable within the componential framework. It is not yet clear whether the componential analysis of intellectual functioning in everyday experience will yield a small enough set of components that can be specifically defined so as to be useful in describing and improving reasoning. It is even less clear whether componential analyses of laboratory reasoning tasks will have much relevance for describing and improving reasoning on real-world reasoning problems. However, in recent work, Sternberg and his colleagues (Sternberg & Caruso, 1985; Sternberg & Wagner, 1986; Wagner & Sternberg, 1985) explored practical intelligence (specifically, tacit knowledge about predictors of on-the-job success, clearly a part of everyday reasoning). Thus, it is quite likely that the issue of how everyday reasoning tasks can be analyzed componentially will receive explicit and detailed attention in the near future.

Specific Rules/Heuristics/Reasoning Schemata Approach

A number of researchers adopt the position that thinking or reasoning amounts to using specific rules or procedures. Good thinking, on this account, is the use of the correct rules, the use of the right number of rules, the correct use of the rules, or some combination of these. This position is most common among researchers who study reasoning within the realm of formal logic, usually propositional logic. The researchers who perhaps best exemplify this position in their work are Martin Braine (1978; Braine et al., 1984), Daniel Osherson (1975), and Lance Rips (1984, 1988).

All three researchers begin with the postulation of a set of inference rules, sanctioned by propositional logic, for which they claim psychological reality. The specific rules proposed vary according to the researchers, but all are stated in terms of propositional variables. If a reasoning problem matches the antecedent of a particular rule in overall form, that rule is invoked and used in the derivation of a conclusion (i.e., the consequent of the rule is then asserted). Rules take the form (antecedent [or premises]) → (consequent [or conclusion]). A specific example from Braine (1978) would be "*premises:* p or q; not p; *conclusion:* q," or, with concrete content, "premises: Either Mondale won or Reagan won; Mondale did not win; conclusion: Reagan won."

Braine (1978) assumed that inference schemata are psychologically elementary and inaccessible to introspection. In addition, he assumed a small set of universal schemata, with few individual differences. Differences in quality of reasoning would most probably result from the failure to use relevant schemata or from the difficulty in putting rules together into a coherent sequence. Likewise, Osherson (1975) postulated that people must possess, in addition to inference rules, "a set of explicit instructions determining the occasions for the use of each [inference rule] in a mental deduction" (p. 85), and he believed that the perceived applicability of an inference rule varies with the context. Rips (1983, 1984) implemented a computer model of reasoning with a set of inference rules (he called them "natural deduction rules") in LISP.

Other researchers have adopted a similar approach to explaining informal reasoning. Cheng and Holyoak (1985), Cheng, Holyoak, Nisbett, and Oliver (1986), and Nisbett, Fong, Lehman, and Cheng (1987), for example, proposed that reasoning about real-life problems relies upon what they called "pragmatic reasoning schemata," that is, generalized context-sensitive sets of rules defined in relation to goals. One example is the permission schema, which can be summarized by four rules: (a) If an action is to be taken, then the precondition must be satisfied; (b) if an action is not to be taken, then the precondition need not be satisfied; (c) if the precondition is satisfied, then the action may be taken; and (d) if the precondition is not satisfied, then the action must not be taken. Such a schema might be activated when reasoning about the following problem: "A patron may consume an alcoholic beverage if she or he was born before January 28, 1967. Susan was born on March 1, 1968. Can Susan consume an alcoholic beverage?". The same schema would not be activated by the problem, "If an executive travels to Boston, she or he takes Northwest Airlines. An executive travels to Montreal. Does she or he take Northwest Airlines?". In the first case, the salient aspects of the problem (including the goal of the rule, viz., to block an underaged patron from drinking) evoke the permission schema, but in the second case, no such evocation occurs, because no salient aspect of the problem suggests the schema.

Pragmatic reasoning schemata, according to Cheng and Holyoak (1985), are context sensitive, are derived through induction from past recurring experiences with relevant situations, and have to do with the goals of the problem. (Notice here the contrast to Braine's [1978] inference schemata, which are context free.) Informal reasoning is based on pragmatic interpretations of problems, and the various schemata keep the inferences that are drawn relevant to the goals of the problem. Good informal reasoning is to be explained in terms of the appropriateness and perhaps the thoroughness with which reasoning problems are mapped onto reasoning schemata. If no pragmatic reasoning schemata can be found, Cheng et al. (1986) allow for the possibility that reasoners would fall back on syntactic (e.g., context free) inference rules, presumably similar to those proposed by Braine (1978) or Rips (1983).

Nisbett, Krantz, Jepson, and Kunda (1983) and Fong, Krantz, and Nisbett (1986) argued that people possess an abstract inferential rule system but only apply the relevant rules when the problem cues them to do so. These authors specifically investigated the question of whether people had a rule corresponding to the law of large numbers. They found that when the problems made the statistical nature of the question clear, people indeed applied a version of this law. Kunda and Nisbett (1986) uncovered similar abilities of people to assess real-world correlations, again under very constrained conditions: (a) when they are highly familiar with the data in question and (b) when the data under consideration are highly capable of being "assigned numerical values to clearly defined units of performance" (p. 198). Good reasoning, for Kunda and Nisbett and others, is again to be explained in terms of appropriate and complete mapping of problems to rules.

When no specific rules seem to apply to a problem, people are assumed to fall back on default rules. Daniel Kahneman and Amos Tversky (1982a, 1982b) postulated a number of such principles or rules that people are believed to use in coming to conclusions in everyday contexts. The function of these heuristics is to "reduce the complex task of assessing probabilities and predicting values to simpler judgmental operations" (Kahneman & Tversky, 1982a, p. 3). Examples of such heuristics include availability (judging frequency or probability of occurrence by how easily examples can be constructed or brought to mind), representativeness (judging the probability that Process B will generate Event A by evaluating the similarity between A and B), or adjustment and anchoring (using an initial value, even an arbitrarily selected one, as a basis from which to estimate a true value). In a number of studies, these researchers demonstrated the (often inappropriate) use of such heuristics, even when heroic efforts are made to control for motivational effects, lapses of memory or attention, or other such factors.

The rules/heuristics approach is useful in explaining content effects in reasoning, that is, why some versions of a problem are easier to solve than formally equivalent versions of the same problems (see Griggs & Cox, 1982; Hoch & Tschirgi, 1985; and Johnson-Laird, Legrenzi, & Sonino-Legrenzi, 1972, for other examples of the phenomenon). Certain contents "naturally" evoke or cue certain rules (Kahneman & Tversky, 1982b), and others do not. However, a detailed description has yet to be offered of just how the evocation process works.

Interestingly, some of the work by philosophers on informal or practical logic seems to fall within the rules/heuristics approach, although in a different way. The goal of such work is to teach people to label and avoid the fallacies that occur in everyday arguments (e.g., unwarranted assumptions, ad hominem attacks), regardless of whether the fallacies are given specific and traditional names. Fallacies are described as patterns of reasoning going wrong in one or more ways (Acock, 1985, Damer, 1980). Damer (1980) classified fallacies into three broad types: (a) those that involve an unacceptable premise; (b) those that involve an irrelevant premise; and (c) those that involve a set of premises, each individually acceptable and relevant, that fail as a set to warrant the conclusion. Strategies for attacking fallacies include the following: reconstructing the argument under consideration so that the error or errors are more clearly evident; challenging an unacceptable premise with a counterexample; or offering an absurd example, one that has the same pattern as the argument under review but that leads to an obviously false conclusion.

Good everyday reasoning, explained in terms of fallacies, is that which is tested against a complete list of fallacies and is found to have no "matches." A reasoner who repeatedly tests her or his (and others') arguments against such a list is claimed to indirectly benefit in the construction of her or his reasoning: "The process of argument evaluation encourages the construction of good arguments by constantly reminding us of the criteria of a good argument" (Damer, 1980, p. 17). That claim, however, has yet to be tested.

Note that the fallacies approach to everyday reasoning hinges on identifying a complete (or at least mostly complete) list of fallacies to ensure good reasoning. This approach falls within the rules/heuristics approach, I think, because one can construe the fallacies as rules that identify flawed patterns of reasoning. Likewise, the practice of comparing an argument to a list of fallacies is best thought of as a heuristic: Unless one can guarantee the completeness of the list, the goodness of the argument is not established absolutely.

Can rules or schemata be taught? The existence of logic and statistics courses in college and university curricula could appear to offer an affirmative answer. However, data are few on how well rules are learned, how long they are retained, and how widely they are applied, especially to everyday reasoning problems.

Work by Fong et al. (1986) and Nisbett et al. (1983, 1987) examined people's understanding of statistical heuristics (e.g., the law of large numbers) in a variety of contexts, including some from real life (e.g., explaining why former sports Rookies-of-the-Year generally perform less well in their second year). They presented subjects with brief training sessions that highlighted the applicability of this heuristic to real-life events. Findings indicated that such sessions improved the subsequent frequency and quality of statistical reasoning. The effects were especially pronounced for subjects who had taken one or more courses in statistics. Cheng et al. (1986), studying conditional reasoning, also reported success in training pragmatic reasoning schemata after brief sessions and also reported that the biggest obstacle was subjects' failure to spontaneously construe a problem in terms of the most appropriate schemata.

With regard to the list of common findings in formal reasoning, the rules/heuristics approach seems to have as its greatest strength the explanation for the intrusion of background knowledge into analytical reasoning problems. By arguing that certain instances of structurally equivalent problems (i.e., contents) cue the appropriateness of the relevant rule more than others do, the rules/heuristics approach can explain differences in performance as a function of content. Regarding premise interpretation problems, a rules/heuristics account might posit specific faulty rules that lead to error but would then incur the burden of specifying where the faulty rules come from and when they apply. This approach could explain the findings regarding people's failure to consider all possible premise interpretations or allowable possibilities with reference to misapplication of a rule or their failure to apply all the relevant rules, presumably again because of the existing cues (or lack thereof) in the problem. The overconfidence phenomenon could be handled in a similar way, although this kind of explanation also would require an explication of what cues are and how they function.

Three major problems confront the rules/heuristics approach. The first is the specification of just what the rules, or heuristics, are. For propositional reasoning, for example, Braine (1978), Osherson (1975), and Rips (1983) proposed different (although overlapping) rule sets, and it is hard to know what measures should be used to judge among them. In terms of other kinds of reasoning (specifically, types of reasoning not fully governed by a traditional system of logic), the problem becomes even more serious, because no constraints have been proposed for what can be a rule. Although Cheng and Holyoak's (1985) examples of pragmatic reasoning schemata are illustrative, for example, it is not clear how one can go about discovering other rule schemata, that is, in what directions their base set can or should be extended and, more importantly, cannot or should not be extended.

The second problem is related to the first: the specification of the origins of the rules. Braine (1978), Osherson (1975), and Rips (1983) all seemed to suggest that inference rules are part of our natural equipment, although Braine et al. (1984) believed that some rules are less available to younger children. For pragmatic reasoning schemata, Cheng and Holyoak (1985) explicitly proposed (and it seems likely that Fong et al. [1986], Kahneman & Tversky [1982a, 1982b], Damer [1980], and Acock [1985] would have implicitly accepted that proposal) that the rules come from inductions from past experience. This account incurs the burden of specifying what governs and constrains that inductive process.

The third issue confronting this approach seems the biggest: the problem of specifying when and how the rules apply. Many of the proposals (especially those of the pragmatic reasoning type) make vague appeals to the salience of the cues within the problem, never tackling the problem of providing a noncircular definition of salience. This seems to push back the problem of explaining reasoning by only one step: People reason by finding the right rule or rules, perhaps, but how do they accomplish this rule selection in the first place? Notice that the fallacies approach also comes under this criticism: There are no algorithms for detecting whether a given fallacy occurs in any argument. This particular problem bears a strong resemblance to the major difficulty with template matching as a theory of perception (Neisser, 1967): explaining how the right cognitive units—be they templates, rules, or heuristics—become activated. In contrast, Rips (1983) had an impressive computer implementation of a reasoning system that automatically selects rules in such a way as to mimic the performance on untutored human reasoners. In this case, however, the constraints on rule selection are imposed by the specificity of the problems, all of them of the formal, propositional reasoning type.

These three problems combine to form a fourth: Until the problems are solved, it is not clear how to extend this approach to reasoning to other kinds of naturally occurring problems of reasoning. Note that in all of the empirical research described here, the problems given to reasoners are self-contained and usually contain abstract or at least arbitrary content. Hence, although much of the work described here purportedly investigates everyday reasoning, I classify it as formal reasoning. In most of the studies cited here, all premises are supplied for the subject by the experimenter. To date, only anecdotal evidence has been offered for the actual use of any of these rules or heuristics in a nonlaboratory setting (see Nisbett & Ross, 1980, for examples). The premises are usually already identified, the amount of irrelevant information is restricted, the number of inferences to be performed is limited to one or a few, there often exist normatively correct answers, and the content of the problems makes little if any contact with issues of real interest or concern to the subjects. Thus, the utility of this approach in describing individual differences in everyday reasoning, or in improving everyday reasoning, remains an open issue.

Mental Models/Search Approach

A third major approach is one that comes from two traditions: Newell and Simon's (1972; Newell, 1981) conception of thinking as searching a problem space and Johnson-Laird's (1982, 1983) idea of mental models. Newell (1981) concisely coined the problem space hypothesis, which held that "the fundamental organizational unit of all human goal-oriented activity is the problem space" (p. 696). By the term *problem space* he meant a set of symbolic structures (states) and operators that take states as input and yield new states as output. To put this in more familiar terms, imagine any given problem. The possibilities allowed by the specification of the problem can be described as lying in some space, with possibilities that are very similar very close together and with possibilities that are very distinct very far apart. What exactly are these possibilities? The answer to that question depends on the specific problem, the domain of the problem, and the particular representation chosen by the problem solver. For chess problems, the possibilities might be the board positions resulting from allowable moves. For categorical syllogisms, the possibilities might be candidate conclusions to the supplied premises. For problems in everyday reasoning and decision making, the possibilities might be the various scenarios generated by considering different options.

Operators, too, vary from domain to domain but always serve a similar function: the construction of states. For chess problems, operators might be the various moves allowable from a given board position. For categorical syllogisms, operators might be the generation of specific possibilities allowed by the premises and suggested by the content of the premises. For problems in everyday reasoning, operators might be changes in the assumptions initially made and the resulting construction of new scenarios based on those changes.

In my view, mental models can be thought of as what Newell (1981) would call states in a problem space. But what is a mental model? Several theorists have used the term in different ways (see Gentner & Stevens, 1983, and Rouse & Morris, 1986, for reviews). In using the term here, I specifically adopt Johnson-Laird's (1983) definition, because he treated mental models within the context of reasoning. Johnson-Laird posited some constraints on what is to count as a mental model: "(a) They [mental models], and the machinery for their construction and interpretation are computable . . . ; (b) they are finite in size . . . ; (c) they are constructed from tokens arranged in a particular structure to represent a state of affairs" (p. 398). He also offered some characterizations that constrain the processes that construct and interpret mental models: (a) When possible, a single model is generated, even if the description is indeterminate; although (b) models can directly represent indeterminacies if their use does not lead to an exponential growth of complexity (Johnson-Laird, 1983, p. 408–409). Finally, he provided a typology of types of mental models, although the list is not meant to be exhaustive: relational, spatial, temporal, kinematic, dynamic, and imaginal. Mental models can be used to represent a true situation, a possible situation, or an imaginal situation.

Given a problem to reason about (either a categorical syllogism or a personal decision), the reasoner is described as generating a number of possibilities (Johnson-Laird [1983] would call these *models*). The generation of a new possibility may in turn raise other problems that in turn call for the generation of other possibilities or models. Presumably, at some point after a possibility is generated, the reasoner assesses its plausibility or credibility by considering the evidence that supports it (Baron, 1985). Finding evidence may present other problems, again calling for the generation of possibilities. At some point, generation of possibilities must halt, and some overall assessment of which possibility (or possibilities) is (or are) the strongest must be made before a conclusion can be drawn.

Reasoning, in this view, amounts to finding a path through a problem space by assembling a sequence of operators that allow legal moves between states (i.e., by constructing a set of one or more different models, consistent with or suggested by the premises). Good reasoning is equated with good searching strategies. Some strategies are weak, domain-independent methods that can vary in terms of effectiveness and efficiency. The generate-and-test strategy, a strategy in which the problem solver or reasoner first constructs or imagines possibilities and then evaluates them to see if they meet relevant criteria, is one example of this type. Typically, when one talks about the generate-and-test strategy, one is thinking about a haphazard or even random order of generation of ideas. The utility of a generate-and-test strategy depends critically on the completeness with which the set of possibilities are generated and also on the rigor of the evaluation metrics. The generate-and-test strategy does not take advantage of domain-specific knowledge.

Other search strategies (depth first, breadth first, progressive deepening, best first; see Winston, 1984, for a review) move closer toward being domain dependent. With experience in a domain, the reasoner develops search control knowledge, which guides the search process toward or away from various regions of the problem space (and possibly helps to create new problem spaces). In the language of models, domain-specific knowledge guides or constrains the construction of one set of models rather than a set of others. Polson and Jeffries (1982) argued that reasoners use domain-specific techniques whenever possible, but when those fail, or if the techniques lack domain-specific knowledge, they fall back on general search techniques, such as the generate-and-test strategy.

Johnson-Laird (1982, 1983) explicitly attempted to use the search framework to understand informal reasoning. He stated the similarities and differences between formal and informal reasoning succinctly: "The rapid implicit inferences of daily life depend on the ability to interpret sentences by constructing mental models of the states of affairs they describe. Deliberate deductions depend on the further ability to search [exhaustively and systematically] for alternative models that violate putative conclusions" (Johnson-Laird, 1982, p. 1). Johnson-Laird explained errors in reasoning in terms of one or more of the following: failure to construct relevant models, failure to search for enough relevant models, failure to search systematically and exhaustively for counterexamples to a putative conclusion, and failure to assess the implications of all the models found in a search. Any of these can (although need not) be directly caused by a shortage of working memory capacity, which in turn hinders the construction and interpretation of mental models.

Perkins and his associates (Perkins, 1986b; Perkins, Bushey, & Faraday, 1986) reported a number of interventions (both naturally occurring—in public and private high schools, colleges, and graduate and professional schools—and specifically designed by the authors, with similar subject populations) to improve reasoning about social and political issues. The kind of training that works best appears to consist of providing content-free "scaffolds," or prompts, that force subjects to generate reasons that run counter to their initial position. Such prompts help lessen what Perkins terms "myside bias", a tendency to generate and consider only those reasons that support one's position in an argument. Other training programs, such as Cognitive Research Trust and the Productive Thinking Program (see Nickerson et al., 1985, for a review), seem to fall between a models/search and a rules/heuristics approach, in that they offer strategic guidelines for exploring an issue (perhaps equivalent to providing rules for constructing models?). However, the effects of these (or any other) programs on everyday reasoning specifically have not been assessed.

This mental models/search approach to reasoning has the following three advantages: First, especially in relation to the previous two approaches, it is more transparently able to describe performance on everyday reasoning tasks as well as on formal ones. This is because it involves the manipulation of representations, and the representations can easily stand for everyday situations. Second, and relatedly, this approach is not constructed in relation to any one particular reasoning task. Therefore, extending it to cover different kinds of reasoning (and even creative thinking or imagining; see Perkins, 1985b) should be relatively easy. Third, this approach accounts for a number of well-replicated findings in the thinking and reasoning literature.

The first finding is that people's general performance in tasks calling for formal logic is mediocre at best. The mental models/search approach handles this fact by pointing out that formal reasoning requires things that everyday inference ordinarily does not: namely, the generation of multiple models and the need to search for models in a systematic and exhaustive manner.

Second, the mental models approach accounts for confirmation bias and belief persistence, both held to be major shortcomings of human thinking (Baron, 1985; Mynatt et al., 1977; Wason, 1977), explaining these both as insufficient and biased search. Further, if it is supposed that a reasoner's confidence is primarily a function of time spent thinking, but a reasoner's search mechanisms are incomplete or biased, this approach also explains the common finding that people generally are overconfident in assessing the efficacy of their own cognition (Fischhoff, 1975, 1982; Lichtenstein et al., 1982): People judge their own efficacy by the time they have spent thinking, but the time they have spent has not been used efficiently.

Third, this approach can account for phenomena found in the problem-solving literature, including set effects and premature closure, and the general tendency to consider too few premise interpretations and allowable possibilities. Set effects would be said to result from a tendency to look in one part of the problem space rather than in others. Premature closure can be accounted for if it is assumed that the quality (i.e., elaborateness) and quantity of models trade off (an assumption made by Perkins, 1981). Thus, the more resources (working memory space, processing time) are invested in any one model, the less likely is the generation of others, leading to premature closure. Although set effects and premature closure are typically associated with problem solving as opposed to reasoning, one could argue that the mental models/search approach explains them in terms of biased searching mechanisms of the sort used in a wide variety of kinds of reasoning.

Finally, the models/search approach accounts for expert–novice differences in terms of the amount of domain-specific search control knowledge and the availability of domain-specific mental models. Note that more sophisticated search strategies reduce the effective problem space, because whole sections of the space are seen as irrelevant or of little utility and thus are left unexplored. This reduction frees processing resources to construct more relevant models or to elaborate the models that are considered. Note, too, that experience in a domain presumably makes available a potentially greater number of mental models and makes the construction of certain models easier and, at the limit, automatic (with enough practice, whole models are recalled instead of constructed). Taken together, these assumptions can explain a classic finding that although grand master and expert chess players sometimes seem to consider about the same number of possible moves, those moves generated by the grand masters are better (de Groot, 1966). The expert has better moves initially available and has an easier time of generating alternatives or further moves and thus is also able to do a more thorough job of searching and testing. Interestingly, other, more extensive work with chess grand masters under other circumstances indicates that more skilled chess players do tend to search more extensively and deeply through the search space, as well as to select higher quality possibilities (Charness, 1981). Charness (1981) explained the individual differences in depth and extensiveness of search in terms of differences in move generation, position assessment, and memory limitations, all of which are presumed to improve with skill.

The mental models/search approach is also not inconsistent with the general finding of content effects in reasoning, although it does not offer a particularly clear explanation either. One could argue that certain contents are easier to encode (whatever that means), therefore taking up fewer working memory resources and leaving more to search and assess other models. One could also assert that certain contents naturally evoke certain models (Markovits, 1984, 1985; Pollard, 1982; Wason & Green, 1984) or that people are more familiar with how to build models with certain contents. Note that this account is quite similar to the account of content effects given by the specific rules/heuristics approach, and thus it is not clear if either approach truly explains content effects.

The models/search approach has greater difficulty explaining the typical finding that untutored reasoners often have problems in encoding the premises. Without extra assumptions, it does not explain why people often misconstrue or alter the meaning of certain premises, particularly those containing negative or marked terms.

The models/search approach also has its share of serious theoretical problems. The first is the lack of definition of what the states in the problem space are. Even if one adopts the idea presented here, that these states are mental models, one still needs to find specific definitional criteria for what constitutes a mental model. A model, as Johnson-Laird (1983) defined it, does not seem distinguishable from what Rumelhart (1980) called a schema, unless perhaps one argues that a model is a schema with all the variables instantiated. Johnson-Laird (1983) provided many examples of models but noted that his list of types is not exhaustive. We need to know what can count as a mental model and what cannot. Rouse and Morris (1986) underscored this point, worrying that without more constraint, the term *mental model* is equated with knowledge, providing very little usefulness to either theorists or investigators.

A related difficulty is the lack of a clear statement about how to look at search as a psychological model. There is a suggestion that the less search one has to do, the better, but not very much is said within psychology about the cases in which search is necessary or just what constitutes search. Polson and Jeffries (1982), studying people's problem solving with problems such as the Towers of Hanoi, argued that novice problem solvers searched more than did experts and created shallower plans. Charness (1981), however, found just the opposite. Little work has been done applying the search metaphor to studies of nonlaboratory thinking and reasoning, however, and it is unclear how well the metaphor fits. Computer scientists, by contrast, posit a number of search techniques that vary in efficiency and effectiveness (Winston, 1984), but few if any psychologists have looked at people's searching behavior using the same framework.

Overall, the biggest shortcoming of the mental models/search approach stems from what I take to be its biggest virtue: its extreme flexibility. Because no constraints have been placed on models or search processes, this approach licenses ad hoc construction of either. Therefore, the theoretical utility of the approach in understanding mental processes awaits the appearance of guidelines that circumscribe the central entities.

General Discussion

A problem common to all three approaches stems from the vagueness of definition of the central terms within each. Although relatively clear examples of components, rules, and models have been offered, the boundaries for the appropriate usage of the terms are undefined. Therefore, it is unclear as to what makes for a component, for example, and even more importantly, what precludes something from being considered a component. Moreover, each of the three approaches fails to specify the processes that use, construct, or manipulate the entities to which the central terms refer. The componential approach, for example, does not specify the conditions under which components are hooked together to be used in a given

Table 2
A Comparison of Three Approaches

Point of comparison	Componential	Rules/heuristics	Models/search
Nature of the fundamental unit	Component (information-processing routine)	Inference rule (rule)	Mental model (representation)
Definition of reasoning	Using one or more particular assemblies of component processes to encode, find, map, or apply a relationship	Following (implicitly or explicitly) a rule, heuristic, or schema to make inferences	Constructing one or more representations consistent with given information to reach and test conclusions
Range of existing reasoning tasks applied to so far	Analogies; series completions; conditional, categorical, and linear syllogisms; mathematical and verbal induction; detective story problems	Conditional, categorical, and linear syllogisms; probabalistic and statistical reasoning problems	Conditional, categorical, and linear syllogisms; open-ended reasoning about social and political issues
Major findings easily explained	Difficulty with negative and marked premise terms; problems in encoding on formal reasoning tasks	Content effects; intrusion of background knowledge	Failure to consider all possible interpretations or consequences of premises; intrusion of background knowledge
Other strengths	Great deal of empirical data that tests theory	Good amount of empirical data that tests theory; has attracted investigators from several fields (philosophy, psychology, artificial intelligence, linguistics)	Specifically designed to be extendable to everyday reasoning
Major shortcomings	Lack of analysis of which components are used in everyday reasoning; lack of definition of the term *component*	Lack of specificity over when and how rules are used; lack of definition of the term *rule*	Lack of empirical data, especially training studies; lack of definition of the terms *model* and *search*

instance of reasoning. Thus, few constraints are placed on the predictions to be derived from the componential approach, reducing empirical testability. Similarly, without specifying the processes by which rules are induced from everyday experience, one cannot evaluate the extent to which the rules/heuristics approach explains a significant proportion of everyday reasoning. The same problem applies to the models/search approach: Some clear and constrained account must be offered of what can and cannot constitute a search process before the approach can be tested.

Without defining central terms and specifying auxiliary processes, no one approach can make a claim to be more empirically testable than any of the others. Therefore, evaluation of the usefulness of each approach must be based on other criteria. One such criterion is the extent to which each approach offers a cogent and compelling explanation of existing findings in the literature, which for the most part describes performance on formal reasoning tasks. A second criterion is the adaptability of each approach to describing and explaining everyday reasoning.

Table 2 summarizes the major points of contrast among the three approaches and highlights the strengths and weaknesses of each. The componential approach has the easiest time accounting for premise encoding and misinterpretation problems. The rules/heuristics approach has as its greatest strength the explanation for content effects and the intrusion of background knowledge. The models/search approach is particularly effective at dealing with failures to consider all premise interpretations or all possibilities allowed by the premises.

How are these conclusions to be integrated? One possibility is to recognize that the three approaches seem to be providing explanations at different levels. The componential approach assumes that the most useful level of description is of the individual cognitive processes used in a task, such as encoding or mapping. As such, the clearest evidence regarding the use of various components comes from performance on self-contained tasks such as analogies and series completion problems. The use of such self-contained problems helps the researcher in the componential tradition isolate and measure parameters associated with the individual cognitive processes, a task that would seem a very difficult one with a richer and therefore noisier instance of reasoning, such as choosing a college or planning an investment strategy. The rules/heuristics approach, in contrast, analyzes reasoning in terms of the specific task-relevant knowledge brought to bear in any particular situation. Therefore, the best research strategy from a rules/heuristics perspective is to try to find structurally similar problems with differing content, such that different stores of knowledge are activated as the content changes. By holding the structure of the problem constant, an important source of task difficulty is presumably controlled, and the effects of central interest—those having to do with domain-specific knowledge—are allowed to emerge. At perhaps a more general level, the mental models/search approach views reasoning from the perspective of a task-independent set of strategies, for example, the strategy of seeking disconfirming evidence to one's hypothesis. Given this approach, it makes sense to vary not only the content but also the structure of reasoning problems under study and to look for similar patterns in performance across a wide-ranging set of tasks.

With regard to everyday reasoning, what role can or should each approach play? Consider first the componential approach. It appears less easily extended to nonlaboratory tasks than the other two. The componential approach requires the presentation to subjects of hundreds of self-contained (and systematically deformed) problems in order to establish parameter values. It therefore seems unlikely that one can test a componential theory of thinking about a personal moral dilemma, because it is hard to tell what would constitute a systematic deformation of the problem. Any real-world personal dilemma typically elicits reasoning over a long (in relation to the laboratory) stretch of time, with episodes of reasoning hard to strictly define. It is therefore difficult to imagine a componential analysis of, say, a person's decision to buy a Honda rather than a Volvo, explicated at a level of detail analogous to the corresponding componential analyses for verbal analogies or categorical syllogisms (Guyote & Sternberg, 1981; Sternberg, 1977a).

Indeed, the advice Sternberg (1986a) provided to enhance real-world reasoning and decision making (e.g., "be sure you consider the full problem" or "don't immediately assume the obvious"; Sternberg, 1986a, pp. 52–53) addresses behavior at the level of strategy (Sternberg would refer here to metacomponents, which may amount to the same thing as strategies) and not to specific performance components. Indeed, the language used here seems more in keeping with the models/search framework than it does with the componential one. Sternberg (1986a) does discuss the role of individual performance components, offering a plethora of exercises in detecting fallacies, working analogies, visual searching, and verbal mapping, but it remains an empirical issue whether extensive practice on these tasks will have a significant impact on everyday reasoning performance specifically.

However, it is possible for the componential approach to play an indirect role in explaining everyday reasoning, by finding laboratory reasoning tasks that predict performance in everyday reasoning. If it can be firmly established that particular instances of laboratory reasoning (e.g., analogies or syllogisms) strongly predict performance on real-world reasoning problems, then it seems likely that processes used by the former tasks are also involved in the latter tasks. Note that the establishment of this connection is far from certain. Some have raised theoretical arguments against the connection (see the arguments by Perkins, 1986a, summarized earlier). And even if one disputes this set of arguments, it remains an empirical question as to whether measures of formal reasoning ability can do a better job predicting everyday reasoning than can general measures of ability, such as intelligence tests. If formal reasoning tasks cannot provide incremental predictive power for real-world reasoning performance, then it is not clear what their utility is to a general theory of reasoning.

The rules/heuristics account appears to operate at a different level of explanation. Rather than focusing on specific cognitive processes used in a task, it concentrates on how the content of the problem cues the knowledge brought to bear on a particular issue. In other words, instead of explaining how one reasons, the rules/heuristics account examines how the subject of one's reasoning effects reasoning performance. This approach assumes that both practice and knowledge allow for the induction of a greater number of rules and for more elaborated rules. This assumption helps provide an account of expert–novice differences in reasoning: Experts, by virtue of their richer knowledge base and extensive experience with problems within a given domain, have a larger and more differentiated set of rules with which to reason.

The extent to which the rules/heuristics approach can be useful in studying everyday reasoning depends upon whether different everyday reasoning contexts are similar enough for common patterns to emerge. It remains to be seen whether real-world reasoning problems, especially those that reasoners care deeply about, are similar in their relevant aspects and occur with enough regularity for rules to develop. Proponents of the rules/heuristics approach will need to delineate the circumstances that allow or facilitate the emergence and the application of rules.

The mental models/search approach appears, at least on the face of it, to be the most easily extendable to everyday reasoning. Because model construction is held to be a frequent everyday activity, the extension of the approach to cover everyday reasoning is straightforward. There is no need to specify what cues construction: It is assumed to be a continually operating process. Just what kinds of models are constructed, and what is required to construct them, however, are less certain. Indeed, part of the very appeal of the models/search approach seems to derive from its major shortcoming: its vagueness, and the lack of extensive empirical testing, which allow it the appearance of great generality. Researchers within the models/search tradition need to specify what is to count as a model and what is to count as search.

Because of its more general level of analysis, the mental models/search approach may be the most useful one for improving everyday reasoning. Consider trying to help someone come to a resolution of a personal dilemma, say, the choice of a career. It is not clear that providing the person with extensive practice with a particular component, such as mapping or encoding, will be of substantial benefit, unless the component being practiced can be shown empirically to play a prominent role in the everyday reasoning problem and to function in similar ways in and out of the laboratory. The rules/heuristics approach seems more helpful. Discussions of, for example, the law of large numbers or the availability heuristic may be of more immediate value to the person choosing a career. However, it is not at all clear that such discussions will benefit the individual later on, when other personal dilemmas not involving that rule are faced.

The models/search approach appears to offer more flexibility than either of the other approaches. It is less crucial within a models/search framework that particular models be taught than it is, for example, within a rules/heuristics framework that particular rules be taught. Instead, reasoners must be taught strategies for generating models in a systematic, if not exhaustive, fashion. What remains to be seen is whether such domain-independent strategies can be identified, taught, and generalized to new situations.

Many (e.g., Baron, 1985; Lipman, 1987; Paul, 1987; Resnick, 1987b) would argue more strongly that cognitive strategies, by themselves, will not improve everyday reasoning. Equally as important as, or perhaps more important than, strategies are dispositions and attitudes about thinking, such as those that help a person accept a prolonged state of doubt, enjoy the process of discovering and challenging ideas, and appreciate the need for procedures that reduce bias. On this view, programs that focus too heavily on specific procedures or rules would seem to be less generally successful with everyday reasoning than would programs that focus on attitudes. Because the mental models/search approach deals more directly with dispositions and attitudes toward thinking (e.g., to avoid impulsivity by searching as systematically and thoroughly as possible, and to be critical of one's favored possibility, by being on guard against biased search) than do the other approaches, it seems more likely to transfer to new situations. In fact, the models/search conception may fit best with a modified rules/heuristics approach, one that emphasizes not rules of inference per se, but rules that apply to inference rules or to the construction and manipulation of models (e.g., assess the problem to see whether statistical principles apply; avoid myside bias). Note that such a combination begins to look like Sternberg's (1984, 1986a) metacomponents, an observation suggesting some implicit convergence among the three approaches.

Naturally, final determination of the strengths and weaknesses of each approach awaits further empirical work on everyday reasoning. Researchers need to begin to define and refine examples of real-world reasoning problems that hold interest and importance for the reasoner. The beginnings of such efforts, reviewed here, are in evidence (see also Rogoff & Lave, 1984, and Sternberg & Wagner, 1986, for examples of other similar efforts). Next, some means of classifying the various examples of everyday reasoning must be found. Until one can describe and measure the nature, frequency, and quality of everyday reasoning, one cannot determine which theoretical approach (or combination of approaches) will be the most beneficial for improving different instances of reasoning in real life. It is hoped in the meanwhile that theorists and investigators of human reasoning will begin to incorporate and address substantive issues of everyday reasoning in their models and in their empirical assessments.

References

Acock, M. (1985). *Informal logic examples and exercises.* Belmont, CA: Wadsworth.

Acredolo, C., & Horobin, K. (1987). Development of relational reasoning and avoidance of premature closure. *Developmental Psychology, 23,* 13–21.

Alexander, P. A., Haensley, P. A., Crimins-Jeanes, M., & White, C. S. (1986). Analogy training: A study of the effects of verbal reasoning. *Journal of Educational Research, 80,* 77–80.

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). San Francisco: Freeman.

Baron, J. (1985). *Rationality and intelligence.* Cambridge, England: Cambridge University Press.

Baron, J. (1988). *Thinking and deciding.* Cambridge, England: Cambridge University Press.

Baron, J. B., & Sternberg, R. J. (Eds.). (1987). *Teaching thinking skills: Theory and practice.* San Francisco: Freeman.

Barratt, B. B. (1975). Training and transfer in combinatorial problem solving: The development of formal reasoning during early adolescence. *Developmental Psychology, 11,* 700–704.

Bartlett, F. (1958). *Thinking.* New York: Basic Books.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85,* 1–21.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 313–371). New York: Academic Press.

Braine, M. D. S., & Rumain, B. (1983). Logical reasoning. In J. Flavell & E. Markman (Eds.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 263–340). New York: Wiley.

Bruner, J. (1957). Going beyond the information given. In H. Gruber (Ed.), *Contemporary approaches to cognition* (pp. 41–69). Cambridge, MA: Harvard University Press.

Bucci, W. (1978). The interpretation of universal affirmative propositions. *Cognition, 6,* 55–77.

Buscaglia, J. (1987). The development of thinking skills within the cultural identity of Latin America. In D. N. Perkins, J. Lockhead, & F. Bishop (Eds.), *Thinking: The Second International Conference* (pp. 141–161). Hillsdale, NJ: Erlbaum.

Case, R. (1974). Structures and strictures: Some functional limits on the course of cognitive growth. *Cognitive Psychology, 6,* 544–573.

Ceci, S. J., & Liker, J. K. (1986). Academic and nonacademic intelligence: An experimental separation. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 119–142). Cambridge, England: Cambridge University Press.

Ceci, S. J., & Liker, J. K. (1988). Stalking the IQ–expertise relation: When the critics go fishing. *Journal of Experimental Psychology: General, 117,* 96–100.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 467–476.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391–416.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293–328.

Clark, H. H. (1969). Linguistic processes in propositional reasoning. *Psychological Review, 76,* 387–404.

Cole, M., & Scribner, S. (1974). *Culture and thought.* New York: Wiley.

Damer, T. E. (1980). *Attacking faulty reasoning* (2nd ed.). Belmont, CA: Wadsworth.

Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly, 28,* 58–64.

de Groot, A. D. (1966). Perception and memory versus thought. In B. Kleinmuntz (Ed.), *Problem solving* (pp. 19–50). New York: Wiley.

Dellarosa, D. (1988). A history of thinking. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 1–18). Cambridge, England: Cambridge University Press.

Detterman, D. K., & Spry, K. M. (1988). Is it smart to play the horses? Comment on "A day at the races: A study of IQ, expertise, and cognitive complexity" (Ceci & Liker, 1986). *Journal of Experimental Psychology: General, 117,* 91–95.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* Boston: Heath.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). San Francisco: Freeman.

Erickson, J. R. (1978). Research on syllogistic reasoning. In R. Revlin & R. E. Mayer (Eds.), *Human reasoning* (pp. 39–50). Washington, DC: V. H. Winston.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

Falmagne, R. (Ed.). (1975). *Reasoning: Representation and process in children and adults.* Hillsdale, NJ: Erlbaum.

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

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 335–351). Cambridge, England: Cambridge University Press.

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

Fredericksen, N. (1986). Human intelligence. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 84–116). Cambridge, England: Cambridge University Press.

Fredericksen, N., Carlson, S., & Ward, W. C. (1984). The place of social intelligence in a taxonomy of cognitive abilities. *Intelligence, 8,* 315–337.

Galotti, K. M., Baron, J., & Sabini, J. P. (1986). Individual differences in syllogistic reasoning: Deduction rules or mental models? *Journal of Experimental Psychology: General, 115,* 16–25.

Galotti, K. M., & Komatsu, L. K. (1989). Correlates of syllogistic reasoning skills in middle childhood and early adolescence. *Journal of Youth and Adolescence, 18,* 85–96.

Geis, M., & Zwicky, A. M. (1971). On invited inference. *Linguistic Inquiry, 2,* 561–566.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Erlbaum.

Glass, A. L., Holyoak, K. J., & Santa, J. L. (1979). *Cognition.* Reading, MA: Addison-Wesley.

Goldman, S. R., & Pellegrino, J. W. (1984). Deductions about induction: Analyses of developmental and individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 149–197). Hillsdale, NJ: Erlbaum.

Grice, H. P. (1975). The logic of conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics III: Speech acts* (pp. 41–58). New York: Academic Press.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73,* 407–420.

Gruber, H. E., & Vonèche, J. J. (1977). *The essential Piaget.* New York: Basic Books.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Guyote, M. J., & Sternberg, R. J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology, 13,* 461–525.

Halpern, D. F. (1984). *Thought and knowledge: An introduction to critical thinking.* Hillsdale, NJ: Erlbaum.

Hawkins, T., 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.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 69,* 366–378.

Henle, M. (1978). Forward. In R. Revlin & R. E. Mayer (Eds.), *Human reasoning* (pp. xiii–xviii). New York: Wiley.

Hoch, S. J., & Tschirgi, J. E. (1985). Logical knowledge and cue redundancy in deductive reasoning. *Memory & Cognition, 13,* 453–462.

Holyoak, K. J., Junn, E. N., & Billman, D. O. (1984). Development of analogical problem-solving skill. *Child Development, 55,* 2042–2055.

Holyoak, K. J., & Nisbett, R. E. (1988). Induction. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 50–91). Cambridge, England: Cambridge University Press.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 255–310). Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1984). Mental speed and the level of analysis. *The Behavioral and Brain Sciences, 7,* 295–296.

Johnson-Laird, P. N. (1982). Thinking as a skill. *Quarterly Journal of Experimental Psychology, 34A,* 1–29.

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

Johnson-Laird, P. N. (1988). A taxonomy of thinking. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 429–458). Cambridge, England: Cambridge University Press.

Johnson-Laird, P. N., Legrenzi, P., & Sonino-Legrenzi, M. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395–400.

Johnson-Laird, P. N., Oakhill, J., & Bull, D. (1986). Children's syllogistic reasoning. *The Quarterly Journal of Experimental Psychology, 38A,* 35–58.

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

Kahneman, D., & Tversky, A. (1982a). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 3–22). Cambridge, England: Cambridge University Press.

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

Keating, D. P. (1980). Thinking processes in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 211–246). New York: Wiley.

Kuhn, D. (1977). Conditional reasoning in children. *Developmental Psychology, 13,* 342–353.

Kunda, Z., & Nisbett, R. E. (1986). The psychometrics of everyday life. *Cognitive Psychology, 18,* 195–224.

Laboratory of Comparative Human Cognition. (1983). Culture and cognitive development. In W. Kessen (Ed.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 1. History, theory and methods* (pp. 295–356). New York: Wiley.

Lave, J., Murtaugh, M., & de la Rocha, O. (1984). The dialectic of arithmetic in grocery shopping. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 67–94). Cambridge, MA: Harvard University Press.

Lesgold, A. (1988). Problem solving. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 188–213). Cambridge, England: Cambridge University Press.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cambridge University Press.

Lipman, M. (1987). Some thoughts on the foundations of reflective education. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 151–161). New York: Freeman.

Markovits, H. (1984). Awareness of the "possible" as a mediator of formal thinking in conditional reasoning problems. *British Journal of Psychology, 75,* 367–376.

Markovits, H. (1985). Incorrect conditional reasoning among adults: Competence or performance? *British Journal of Psychology, 76,* 241–247.

Matlin, M. (1983). *Cognition.* New York: Holt, Rinehart & Winston.

McDermott, D. (1981). Artificial intelligence meets natural stupidity. In J. Haugeland (Ed.), *Mind design* (pp. 143–160). Montgomery, VT: Bradford Books.

Moshman, D., & Franks, B. A. (1986). Development of the concept of inferential validity. *Child Development, 57,* 153–165.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking* (pp. 315–325). Cambridge, England: Cambridge University Press.

Neimark, E. D. (1987). *Adventures in thinking*. New York: Harcourt Brace Jovanovich.

Neimark, E. D., & Chapman, R. H. (1975). Development of the comprehension of logical quantifiers. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 135–152). Hillsdale, NJ: Erlbaum.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Neisser, U. (1983). Components of intelligence or steps in a routine? *Cognition, 15,* 189–197.

Newell, A. (1981). Reasoning, problem solving, and decision processes: The problem space as a fundamental category. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 693–718). Hillsdale, NJ: Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nickerson, R. S., Perkins, D. N., & Smith, E. E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science, 238,* 625–631.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Olsen, D. R. (1986). Intelligence and literacy: The relationships between intelligence and technologies of representing and communication. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 338–360). Cambridge, England: Cambridge University Press.

Osherson, D. N. (1975). Logic and models of logical thinking. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 81–92). Hillsdale, NJ: Erlbaum.

Overton, W. F., Ward, S. C., Novek, I. A., Black, J., & O'Brien, D. P. (1987). Form and content in the development of deductive reasoning. *Developmental Psychology, 23,* 22–30.

Paul, R. W. (1987). Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 127–148). New York: Freeman.

Pellegrino, J. W., & Glaser, R. (1980). Components of inductive reasoning. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), *Aptitude, learning and instruction* (Vol. 1, pp. 177–217). Hillsdale, NJ: Erlbaum.

Pellegrino, J. W., & Lyon, D. (1979). The components of a componential analysis. *Intelligence, 3,* 169–186.

Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Perkins, D. N. (1985a). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77,* 562–571.

Perkins, D. N. (1985b). Reasoning as imagination. *Interchange, 16,* 14–26.

Perkins, D. N. (1986a). *Knowledge as design*. Hillsdale, NJ: Erlbaum.

Perkins, D. N. (1986b, April). *Reasoning as it is and could be: An empirical perspective*. Paper presented at the AREA Conference, San Francisco, CA.

Perkins, D. N., Allen, R., & Hefner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell (Ed.), *Thinking: The expanding frontier* (pp. 177–189). Philadelphia, PA: Franklin Institute.

Perkins, D. N., Bushey, B., & Faraday, M. (1986). *Learning to reason* (Contract No. NIE-G-83-0028). Washington, DC: National Institute of Education.

Pollard, P. (1982). Some possible effects of availability. *Cognition, 12,* 65–96.

Polson, P., & Jeffries, R. (1982). Problem solving and search and understanding. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 367–411). Hillsdale, NJ: Erlbaum.

Raths, L. E., Wasserman, S., Jonas, A., & Rothstein, A. (1986). *Teaching for thinking* (2nd ed.). New York: Teachers College Press.

Resnick, L. B. (1987a). Learning in and out of school. *Educational Researcher, 16,* 13–20.

Resnick, L. B. (1987b). *Education and learning to think*. Washington, DC: National Academy Press.

Revlis, R. (1975). Syllogistic reasoning: Logical decisions from a complex data base. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 93–135). Hillsdale, NJ: Erlbaum.

Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90,* 38–71.

Rips, L. J. (1984). Reasoning as a central intellective ability. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 105–149). Hillsdale, NJ: Erlbaum.

Rips, L. J. (1988). Deduction. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 116–152). Cambridge, England: Cambridge University Press.

Roberge, J. J. (1970). A study of children's abilities to reason with basic principles of deductive reasoning. *American Educational Research Journal, 7,* 583–596.

Roberge, J. J., & Craven, P. A. (1982). Developmental relationships between reading comprehension and deductive reasoning. *The Journal of General Psychology, 107,* 99–105.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100,* 349–363.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Scribner, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking* (pp. 483–500). Cambridge, England: Cambridge University Press.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 13–30). Cambridge, England: Cambridge University Press.

Shapiro, B. J., & O'Brien, T. C. (1970). Logical thinking in children ages six through thirteen. *Child Development, 41,* 823–829.

Smith, L. B., Sera, M., & Gattuso, B. (1988). The development of thinking. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 366–391). Cambridge, England: Cambridge University Press.

Snow, R. E., Kyllonen, P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–103). Hillsdale, NJ: Erlbaum.

Staudenmayer, H., & Bourne, L. E., Jr. (1977). Learning to interpret conditional sentences: A developmental study. *Developmental Psychology, 13,* 616–623.

Sternberg, R. J. (1977a). Component processes in analogical reasoning. *Psychological Review, 84,* 353–378.

Sternberg, R. J. (1977b). *Intelligence, information-processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1982a). Reasoning, problem solving, and intelligence.

In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). Cambridge, England: Cambridge University Press.

Sternberg, R. J. (1982b). A componential approach to intellectual development. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 413–463). Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition, 15,* 1–48.

Sternberg, R. J. (1984). Toward a triarchic theory of human intelligence. *The Behavioral and Brain Sciences, 7,* 269–315.

Sternberg, R. J. (1986a). *Intelligence applied.* New York: Harcourt Brace Jovanovich.

Sternberg, R. J. (1986b). Toward a unified theory of human reasoning. *Intelligence, 10,* 281–314.

Sternberg, R. J., & Caruso, D. R. (1985). Practical modes of knowing. In E. Eisner (Ed.), *Learning and teaching the ways of knowing* (pp. 133–158). Chicago: University of Chicago Press.

Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General, 112,* 80–116.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology, 27,* 195–232.

Sternberg, R. J., & Smith, E. E. (Eds.). (1988). *The psychology of human thought.* Cambridge, England: Cambridge University Press.

Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence.* Cambridge, England: Cambridge University Press.

Taplin, J. E., Staudenmayer, H., & Taddonio, J. L. (1974). Developmental changes in conditional reasoning: Linguistic or logical? *Journal of Experimental Child Psychology, 17,* 360–373.

Vye, N. J., Delclos, V. R., Burns, S., & Bransford, J. D. (1988). In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 337–365). Cambridge, England: Cambridge University Press.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49,* 436–458.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 51–83). Cambridge, England: Cambridge University Press.

Wason, P. C. (1977). On the failure to eliminate hypotheses—A second look. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking* (pp. 89–97). Cambridge, England: Cambridge University Press.

Wason, P. C., & Green, D. W. (1984). Reasoning and mental representation. *Quarterly Journal of Experimental Psychology, 36A,* 597–610.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* Cambridge, MA: Harvard University Press.

Wertheimer, M. (1945). *Productive thinking.* New York: Harken.

Wertsch, J., Minick, N., & Arns, F. (1984). The creation of context in joint problem-solving. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 151–171). Cambridge, MA: Harvard University Press.

Winston, P. H. (1984). *Artificial intelligence* (2nd ed.). Reading, MA: Addison-Wesley.

Received December 7, 1987
Revision received September 9, 1988
Accepted October 17, 1988 ■