

# The interaction between reasoning and decision making: an introduction

Philip N. Johnson-Laird, Eldar Shafir

*Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544, USA*

In order to decide, judge;  
in order to judge, reason;  
in order to reason, decide (what to reason about).

Reasoning and decision making are high-level cognitive skills that have been under intensive investigation by psychologists and philosophers, among others, for the last thirty years. But, methods and theories have developed separately in the two fields so that distinct traditions have grown with little to say to one another. The aim of this special issue of *Cognition* is to encourage and help workers in these two traditions to understand one another's research and to reflect and enhance some recent signs of "crosstalk" between them.

It seemed to us high time to consider the growing interactions between reasoning and decision making for at least three reasons. First, the two abilities are often interwoven in real life, at least according to the common-sense view epitomized in the maxim at the head of this introduction. Second, their study has led to striking parallels in the conclusions that investigators have reached and to the recent signs of crosstalk. Third, the two fields have important lessons for each other. We will explore each of these reasons before we introduce the papers in this special issue. We begin with a sketch of the everyday assumptions that relate reasoning and decision making – a background that we will defend, though it is sometimes disparaged by philosophers as "folk psychology".

## **The "folk" psychology of reasoning and decision making**

Human beings have desires and needs, and they use their knowledge to decide what to do and to infer how best to achieve their goals. They reason in order to make decisions and to justify them both to themselves and others; they reason in

*Correspondence to:* P.N. Johnson-Laird, e-mail: phil@clarity.princeton.edu or to E. Shafir, e-mail: eldar@clarity.princeton.edu; fax: (609) 258 1113.

order to determine the consequences of their beliefs and of their hypothetical actions; they reason to work out plans of action. They make decisions about what values to treat as paramount; they make decisions about what actions to take; and they make decisions about what information to base their reasoning on. Hence, there is an interdependence between reasoning and decision making. They are, as computer scientists say, mutually recursive.

This account of reasons and decisions is part of “folk psychology”, that is, the view that most individuals in our culture accept about mental life. Philosophers and others have a variety of views about folk psychology. Thus, at one extreme, Ramsey (1926/1990) brought together three of its fundamental concepts – truth, probability, and value – in a profound synthesis. At another extreme, however, certain philosophers reject folk psychology wholesale. Terms such as “belief” and “desire”, they say, do not refer to any substantive entities. They are theoretical terms much like “phlogiston” or “ether”, which will be replaced once science has succeeded in determining the real underpinnings of behavior. People may use folk psychology to predict their own and other individuals’ actions, but, say the critics, the theory itself is incorrigible – it is not responsive to evidence; it is not a scientific theory. The proper study of the mind is instead brain, nerve, and synapse (e.g., Churchland, 1984; Stich, 1983). At the heart of these critiques seems to lie a common philosophical skepticism about ‘intentional objects’, that is, mental representations that have content and that are about something.

Students of reasoning and decision making do not generally regard assumptions about beliefs, desires, and values with total skepticism, but they are suspicious of one component of folk psychology; the myth of perfect awareness. Unlike cognitive and social psychologists, most people believe that their conscious feelings and judgments control their actions. The claim is a legacy of the Cartesian identification of the mind with consciousness. Nothing is easier according to this principle than to know the contents of your mind: introspection will deliver them to you as part of your immediate experience. But to what extent can you be aware of the causes of your behavior? There is probably no way to answer this question for any particular action. That is the problem that historians face. But the psychological study of reasoning and decision making has answered the question for replicable classes of action. Individuals are often not aware of how they reason, having at best only glimpses of the process. They are aware of the results, not the mechanism. What they say about their reasoning does not tally with its real nature (e.g., Wason & Evans, 1975). Likewise, as Nisbett and Ross (1980) have argued, people often are not aware of the real basis of their decisions. For example, Nisbett and Wilson (1977) asked shoppers in a mall to rate the quality of four nightgowns. Most subjects favored whichever was the last gown that they had examined, while attributing their preference to some property of the nightgown itself, rather than its position in the sequence. In a similar vein, psychologists have demonstrated that individuals’ simplistic theories about the

social roots of behavior pervade their attributions of causes for others' actions and emotions.

Yet, the imperfections of introspection do not imply that all aspects of folk psychology are illusory. Individuals do have access to goals, feelings, and values. Introspection makes available to them what they are thinking, and in this way, as Ericsson and Simon (1980) contend, it can provide clues to the underlying process. People regularly ascribe truth values to assertions. As long as this practice is not entirely suspect, then it is hard to see how there could be anything dubious about the idea that an individual holds an assertion to be true (or false). And thus the concept of belief remains in the domain of respectable discourse. It is true, nonetheless, that introspection is not a direct route to understanding mental processes, and, as far as we know, there is no direct route. That is why psychologists studying reasoning and decision making are committed to experimental investigations rather than to introspection and *a priori* analysis (also known as 'indoor ornithology').

### **Normative theories are not psychological theories**

Normative theories have been proposed both for reasoning and for decision making. They are intended to specify what count as rational inferences and rational decisions. In the case of reasoning, the normative theory is formal logic, and many theorists have argued that it does indeed provide an account of human deductive competence (e.g., Beth & Piaget, 1966; Cherniak, 1986; Henle, 1962; Macnamara, 1986). There are, however, three difficulties with this proposal. First, which logic gives the normative account? Logic is not a monolithic enterprise, particularly if it is to embrace reasoning with modal terms such as "possibly" and "necessarily": there are probably an infinite number of distinct modal logics (Hughes & Cresswell, 1968). Perhaps none of them is relevant to human competence, because an assertion, such as:

It is raining, and possibly it is not raining.

is not a self-contradiction in any of these logics, as it seems to be in everyday language, which distinguishes between this assertion and the following "counterfactual" one:

It is raining, but it might not have been raining.

Second, logic gives no account of which valid conclusion to draw. From any set of premises, there are always infinitely many conclusions that follow validly; that is, they must be true given that the premises are true. Most of these conclusions are

trivial, and it would hardly be rational to infer them. For example, given the following premises:

If it is raining, then there is no need to water the plants.

It is raining.

a sensible, rational, conclusion is:

There is no need to water the plants.

But any of the following conclusions are valid too:

It is raining or it is not raining.

It is raining or it is snowing.

It is raining or cats and dogs are falling out of the sky.

and so on.

Logic alone cannot determine which of the infinitely many valid conclusions are sensible to draw. Strangely, the theories that adopt logic as their account of deductive competence have almost universally overlooked this problem. Human reasoners seem naturally to infer conclusions that are parsimonious, that make information explicit that was not stated as such in the premises, and that do not throw information away by adding disjunctive alternatives (see Johnson-Laird & Byrne, 1991).

Whatever the definition of competence, human reasoners do make deductive errors. Yet, there have been many attempts to save the hypothesis that they are logically impeccable. The idea was first formulated by philosophers, perhaps on the grounds that humans are made in God's image. More recently, Henle (1962) has argued that apparent mistakes in reasoning do not impugn the underlying mechanism: individuals *are* reasoning validly, but they have forgotten a premise, reinterpreted a premise, or imported some additional premise. Similarly, Cohen (1981) has argued that the underlying competence cannot be at fault: there is merely a malfunction in the mechanism. The evidence suggests otherwise. Reasoners make errors in cases where there are no grounds for supposing that they have forgotten the premises, distorted them, or added new premises. Their errors are systematic and predictable, and they will even concede that they have erred. The phenomena do not imply that human beings are irredeemably irrational – after all, they are sensible enough to recognize their shortcomings and to devise formal logic as a device for testing the validity of their inferences. Systematic errors in deduction are compatible with the notion that deductive competence rests on the semantic principle of validity. According to this notion, an inference is valid if its conclusion must be true given that its premises are true. Human beings appear to grasp this idea (at least tacitly), but they do not always

have the mental resources to pursue it properly, and so they fall into error. This idea lies at the heart of the mental model theory of reasoning, which is outlined in this issue (see the paper by Legrenzi, Girotto, & Johnson-Laird).

In the case of decision making, the classical normative account is expected utility theory (EUT), wherein the value of an alternative consists of the sum of the utilities of its outcomes, each weighted by its probability. Modern EUT was developed by von Neumann and Morgenstern (1947), who showed that if individuals' preferences satisfy a number of simple axioms, then their behavior can be described and justified as maximizing expected utility. Besides some more technical points, the axiomatic analysis reflects the following substantive assumptions about decisions (for further discussion, see Tversky & Kahneman, 1986, and references therein; in what follows we adopt their terminology, although these notions appear in a number of different guises elsewhere):

*Cancellation:* Any state of the world that yields the same outcome regardless of one's choice should be "cancelled", or eliminated from further consideration. Thus, choice between options should depend only on those states in which they yield different outcomes.

*Transitivity:* For any alternatives  $x$ ,  $y$ , and  $z$ , if  $x$  is preferred to  $y$ , and  $y$  is preferred to  $z$ , then  $x$  must be preferred to  $z$ .

*Dominance:* If one option is better than another in one state of the world and at least as good in all other states, then the dominant option should be chosen.

*Invariance:* Different but logically identical representations of the same choice problem, or different methods of eliciting a choice, should yield the same preferences.

Just as the study of judgment has shown that humans are not intuitive Bayesian statisticians (see the papers in this issue by Klayman & Brown and by Smith, Shafir, & Osherson), so too studies of decision making have shown that people are not intuitive utility maximizers. Research in behavioral decision theory has documented systematic and predictable violations of the above assumptions. Perhaps the most basic assumption is that of invariance, which requires strategically equivalent methods of elicitation and logically identical descriptions of the options to yield the same preferences. Violations of invariance have now been documented in numerous domains, in hypothetical as well as real world situations, with both high and low stakes, and both with and without monetary incentives. As illustrated by Tversky and Kahneman (1986), violations of

invariance themselves suffice to generate violations of cancellation and dominance. Some violations of transitivity have also been documented (Tversky, 1969), and are most likely to arise in decisions based on multiple criteria. They can also occur in collective behavior, as illustrated by well-known paradoxes in social choice (e.g., Arrow, 1963).

Certain assumptions of rational choice have had more normative appeal than others. Thus, while the assumption of invariance is indispensable for a normative account, the cancellation and transitivity conditions have been challenged by some theorists. The notion that humans abide by normative principles dies hard, especially amongst economists and philosophers (see Hogarth & Reder, 1986, for a good discussion.) As a consequence of this tension, several theories have been proposed that retain some of the more normatively appealing principles, like invariance, while relaxing others, such as cancellation and transitivity (for reviews, see Camerer, 1993; Machina, 1982; Schoemaker, 1982). Their apparent aim is to tailor a normative account to fit observed behavior. Naturally, as evidence continues to accumulate documenting the violation of even the most essential principles, decision theorists, like students of human reasoning, will have to accept the irreconcilability of the normative and the descriptive accounts.

In sum, the major psychological discovery about both reasoning and decision making is that normative theory and psychological facts pass each other by. People are not intuitive logicians, intuitive statisticians, or intuitive rational decision theorists. Instead, the precise character of their thoughts and decisions is the outcome of complex and unobservable mental processes, the nature of which researchers in both these areas of inquiry are trying to elucidate.

### **Some lessons from the study of reasoning and decision making**

A significant lesson from the study of reasoning is that theories of mental processes should be computable, that is, sufficiently explicit that, in principle, they could be modeled in computer programs. Indeed, many researchers have implemented their theories in computer programs (e.g., Braine & O'Brien, 1991; Johnson-Laird & Byrne, 1991; Rips, 1983). The practice helps to ensure that theorists are not taking too much for granted and that their theories are not vague, incoherent, or, like mystical insights, only properly understood by their proponents. The effort to develop a computable theory concentrates the theorist's mind on the nature of underlying mental processes; it alerts the theorist to the design of mental architecture (Newell, 1990) and to the puzzle of how intellectual strategies might develop. Similarly, since programs should not be equated with theories, there can be a dialectical improvement in the theories: what a program

does is not as important as the effects its development may have on the theorist's thinking.

A significant lesson from the study of decision making is the need for experiments to examine realistic problems. It is of limited interest to investigate text-book examples if they are so artificial and so remote from the exigencies of daily life that subjects are likely to adopt strategies that are otherwise alien to their everyday thinking. The result will be a burgeoning literature that is largely irrelevant to the real nature of thinking. In contrast, students of decision making have investigated, among others, the decisions of doctors about treatments, of gamblers in casinos, and of customers in shopping malls. Similarly, many of the laboratory studies, while hypothetical in nature, show an enhanced sensitivity to the verisimilitude of their tasks.

Another lesson from the study of decision making is the three-fold distinction that theorists draw among normative, descriptive and prescriptive accounts (e.g., Bell, Raiffa, & Tversky, 1988). The customary distinction is between the normative (how one ought to proceed) and the descriptive (how one in fact proceeds). But, if individuals are to be helped to improve their performance in daily life, it is unrealistic to hope that this may be achieved simply by teaching them decision theory. Recommendations need to go beyond normative and descriptive accounts and to advocate measures that actually help people make better decisions. Such prescriptive measures are likely not to fall neatly into either the normative or descriptive categories. If a decision problem is affected by the way in which it is framed, then a combination of the alternative frames into a single description may not overcome the effect: one frame may still be more salient than the other (cf. McNeil, Pauker, & Tversky, 1988). At that point, the best that decision makers can do is to examine their preferences from the different perspectives and, aware of the normative–descriptive tension, reason about the best way of reaching a decision. The development of effective prescriptions to overcome such difficulties calls for further thinking and analysis on the part of both lay people and researchers.

To a first approximation, these are the lessons that the study of reasoning and decision making have for one another: investigators of reasoning often use artificial materials or tasks remote from daily life, and seldom recognize the need for special prescriptive measures (but compare Bauer & Johnson-Laird, in press); investigators of decision making are often more concerned with what the mind is doing than with how it is doing it, that is, with specifying a computable account of the relevant mental processes (but compare Payne, Bettman, Coupey, & Johnson, 1992). We hope this special issue of *Cognition* will help workers in the two areas understand one another's research and share their insights. That must be the first step towards the integration of the two areas into a larger explanatory framework.

### **An introduction to the papers**

The papers in this issue range over a variety of sorts of inferences and decisions. They express diverse points of view, propose different theories, and take different positions on the topic of human rationality. What they have in common, we believe, is a commitment to the idea that students of reasoning and decision making both stand to benefit from a closer interaction and dialogue with one another, and from a more interactive sharing of theories, methodologies, and findings.

Shafir, Simonson, and Tversky argue that a consideration of the reasons that enter into people's thinking about a decision problem may illuminate certain aspects of reflective choice that remain counterintuitive from the perspective of the classical theory.

Legrenzi, Girotto, and Johnson-Laird assume that reasoning and decision making both depend on the construction of mental models. They demonstrate that when individuals make deductive inferences or seek information in order to make decisions, they focus on what is explicitly represented in their models to the neglect of other possible information.

Smith, Shafir, and Osherson examine inductive inferences based on category membership and the heuristics that influence judgments of argument strength, that is, the probability that a conclusion is true given that the premises are true. They show that with unfamiliar predicates what matters most is the similarity between the premise and conclusion categories, whereas with familiar predicates probability judgments are based on both the similarity relations and the plausibility of the premises and conclusion.

Klayman and Brown consider what can be done to improve human judgment and reasoning. They show that one stratagem – changing the environment, that is, the way in which information is presented – can reduce error in a task based on medical diagnosis, moving subjects away from focusing on typical features to those that are genuinely diagnostic.

Pennington and Hastie analyze the reasoning of members of juries trying to reach a decision, and show how they rely on explanations that, in turn, are based on inferences that lead to a causal model of the events.

Evans, Over, and Manktelow take up the question of rationality. They distinguish between reasoning to achieve goals and reasoning as a logical process, and argue that the latter is not the basis for the former. They conclude that both decision theory and formal logic are inadequate models for assessing human competence.

Inevitably, much excellent work exists for which there was no room, but we have been fortunate that nearly everyone whom we invited has written an article for this special issue. We thank the authors for their papers and their patience, and the referees for their criticisms and suggestions. We also thank Jacques

Mehler, the editor of the journal, and Susana Franck, the editorial associate, for their help, advice, and encouragement.

## References

- Arrow, J. (1963). *Social choice and individual values*, 2nd edn. New York: Wiley.
- Bauer, M.I., & Johnson-Laird, P.N. (in press). How diagrams can improve reasoning. *Psychological Science*.
- Bell, D.E., Raiffa, H., & Tversky, A. (1988). Descriptive, normative, and prescriptive interactions in decision making. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: descriptive, normative, and prescriptive interactions*. Cambridge, UK: Cambridge University Press.
- Beth, E.W., & Piaget, J. (1966) *Mathematical epistemology and psychology*. Dordrecht: Reidel.
- Braine, M.D.S., & O'Brien, D.P.O. (1991). A theory of "if": a lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 98, 182–203.
- Camerer, C.F. (1993). Recent tests of generalizations of expected utility theory. In W. Edwards (Ed.), *Utility theories: measurements and applications*. Hingham, MA: Kluwer.
- Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.
- Churchland, P.M. (1984). *Matter and consciousness*. Cambridge, MA: MIT Press.
- Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317–370.
- Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87, 215–251.
- Henle, M. (1962). The relation between logic and thinking. *Psychological Review*, 69, 366–378.
- Hogarth, R.M., & Reder, M.W. (Eds.) (1986). The behavioral foundations of economic theory. *Journal of Business*, 59 (4), Part 2.
- Hughes, G.E., & Cresswell, M.J. (1968). *An introduction to modal logic*. London: Methuen.
- Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.
- Machina, M.J. (1982). "Expected utility" analysis without the independence axiom. *Econometrica*, 50 277–323.
- Macnamara, J. (1986). *A border dispute: the place of logic in psychology*. Cambridge, MA: Bradford Books/MIT Press.
- McNeil, B.J., Pauker, S.G., & Tversky, A. (1988). On the framing of medical decisions. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: descriptive, normative, and prescriptive interactions*. Cambridge, UK: Cambridge University Press.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Nisbett, R.E., & Ross, L. (1980) *Human inference: strategies and shortcomings of social judgement*. Englewood Cliffs, NJ: Prentice Hall.
- Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Payne, J.W., Bettman, J.R., Coupey, E., & Johnson, E.J. (1992). A constructive process view of decision making: multiple strategies in judgment and choice. *Acta Psychologica*, 80, 107–141.
- Ramsey, F.P. (1990). Truth and probability. In Mellor, D.H. (Ed.), *Philosophical papers: F.P. Ramsey*. Cambridge, UK: Cambridge University Press. (Original work published 1926).
- Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90, 38–71.
- Schoemaker, P.J.H. (1982). The expected utility model: its variants, purposes, evidence, and limitations. *Journal of Economic Literature*, 20, 529–563.
- Stich, S. (1983). *From folk psychology to cognitive science*. Cambridge, MA: MIT Press.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59 (4), part 2, S251–S278.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton: Princeton University Press.
- Wason, P.C., & Evans, J.St.B.T (1975) Dual processes in reasoning? *Cognition*, 3, 141–154.