MENTAL MODELS: Descriptive and Normative Aspects

James H. Fetzer
Department of Philosophy
University of Minnesota
Duluth, MN 55812

ABSTRACT: A theory of "mental models" has been elaborated by Philip Johnson-Laird (1983) and applied to deductive reasoning by Johnson-Laird and Ruth M. Byrne (1991). They have now extended their approach to conditionals as a theory of meaning, of pragmatics and of inference (forthcoming). A critique of their latest efforts has been advanced by David Over and Jonathan St. B. T. Evans (forthcoming), which focuses on properly understanding conditionals, especially relative to probability. My purpose here is to assess Over and Evan's criticism and add a few of my own, in an effort to clarify the respective contributions of logical, subjective, frequency and propensity conceptions of probability to the pragmatics of ordinary reasoning and the standing of Lewis' Principal Principle. It should become quite apparent why Johnson-Laird and Byrne's theory is neither the only nor the best within this domain.

KEYWORDS: mental models, ordinary reasoning, subjective probability, frequency frequency probability, propensity probability, logical probability, the Ramsey test, material vs. subjunctive conditionals, hypothetical reasoning, The Principal Principle

A theory of "mental models" has been elaborated by Philip Johnson-Laird (1983) and applied to deductive reasoning by Johnson-Laird and Ruth M. Byrne (1991). They have now elaborated their approach for conditionals generally as a theory of meaning, of pragmatics and of inference (forthcoming). A critical evaluation of their latest efforts has been advanced by David Over and Jonathan St. B. T. Evans (forthcoming), which focuses on the proper understanding of conditionals, especially in relation to probability. Since I have addressed (what I take to be) the strengths and weaknesses of mental model theory in its earlier incarnations in several other places (Fetzer (1993), (1999a) and (1999b)), I shall direct most of my attention to Over and Evans' critique. It should become apparent that Johnson-Laird and Byrne's theory is neither the only nor the best within this domain. My approach, therefore, shall be as follows. First, I shall summarize a few of the fundamental problems that I find in Johnson-Laird and Byrne's theory of mental models. Second, I shall review and endorse many of the criticisms Over and Evans direct against their most recent work on conditionals. Third, I shall extend what Over and Evans have had to say by distinguishing between subjective and logical probabilities, on the one hand, and relative frequencies and causal propensities, on the other. Fourth, I shall discuss the Principal Principle that David Lewis (1983) has advanced as the foundation for relating subjective and objective probabilities, including a new and improved formulation (Lewis 1994). Fifth, I shall discuss other kinds of mental models derived from the field known as cognitive geography that display alternative conceptions of the terrain thereby subsumed.

1. Preliminary considerations.

The notion of a model is sufficiently ambiguous that any mental activity where something stands for something else in some respect or other might properly qualify. What matters is the meaning attached to that phrase by specific students in a specific context. Johnson-Laird and Byrne (henceforth, JL and B) maintain that subjects who are untutored in logic reason in three stages of thought involving comprehension of the premises (based upon background knowledge and linguistic understanding), their parsimonious redescription (making previously implicit content explicit, for example), and validation (as a search for alternative models in which the premises are true but the conclusion false). The conclusion then "follows from its premises" when there are no counterexamples where the premises are true and the conclusion simultaneously false. What they thereby introduce is a notion of valid-for-z, where "z" is replaced by the name or description of a particular person. The mental powers of different persons can differ enormously, ranging as they do over newly born infants, babes-in-arms, and prepubescent children, not to mention the aged, the infirm, and the mentally bewildered. Indeed, what is valid-for-z at time t1 may or may not be valid-for-z at time t2, any more than what is valid-for-z1 may not be valid-for-z2 at the same time t. These notions are complemented by counterparts for possible-for-z, probable-for-z, and the like, where an outcome is possible-for-z if z has at least one mental model in which the conclusion is true when the premises are true and probable-for-z when z has more mental models in which the conclusion is true than z has mental models in which the conclusion is false. z's personal performance can of course diverge arbitrarily from objective competence.

2. Logic vs. psychology.

Among the most important differences between logic and psychology with regard to arguments and inference is that one is normative while the other is descriptive. Logic attempts to codify the principles and the patterns of inductive and of deductive reasoning when that reasoning is proper or valid, even apart from questions of truth. An inductive argument that accurately represents the degree of evidential support provided by a given set of premises is said to be "proper", just as a counterpart deductive argument is said to be "valid". When a proper inductive argument has true premises, it qualifies as "correct"; and when a valid deductive argument has true premises, it qualifies as "sound" (Fetzer 1981, p. 177). Apart from proofs of logical truths, the arguments studied by logicians are hypothetical and concern what follows from or receives support from what, independent of the truth of the premises that are given, whether the context is inductive or is deductive. Johnson-Laird and Byrne segregate thought processes into five categories, namely: associative, creative, calculating, deductive, and inductive (1991, p. 2). Those they regard as species of reasoning specifically as opposed to modes of thinking generally are therefore calculation, deduction, and induction. JL and B characterize "calculation" as "the routine application of a procedure known by heart, as in mental arithmetic." Mathematics, even including (what is known as) the principle of mathematical induction, is deductive in the sense that, for valid mathematical arguments, their conclusions follow from their premises, necessarily. Given "2" and "2" as natural number inputs, the operation of addition ("+") in abstract contexts, and the standard axioms of arithmetic, it cannot be false that the sum of 2 and 2 equals 4. There appears to

be no foundation for separating out "calculation" as a species of reasoning apart from induction and deduction. Deduction includes calculation.

3. The "principle of truth".

Others may be as puzzled as I am by the "principle of truth" that Johnson-Laird and Byrne advance. They deny that mental model theory relies on truth tables, which is a standard semantical method for the study of logic, while maintaining that "reasoners normally represent only what is true. In this way, they minimize the load on their short-term memory" (1999, p. 112). But this is very odd on at least two counts. The first is that reasoners are typically reasoning about matters where they do not yet know what is true and what is false: that is why they are reasoning! Their premises are taken as true hypothetically, not historically. Consider: "If we invade Iraq, then that will open the gates of hell in the Middle East. We are going to invade Iraq. Therefore, we are going to open the gates of hell in the Middle East." There would be no point to reasoning of this kind if had to already know the truth of our premises. It would be too late to matter. JL and B maintain that their "principle of truth" is subtle because it operates at two levels: "First, reasoners represent only true possibilities. Second, within the true possibilities, they represent only those literal propositions (affirmative or negative) in the premises that are true" (1999, p. 112). The use of the term "true possibility", of course, introduces profound ambiguity. Relative to the example at hand, there are four different possibilities, yielding a truth table that displays four possible combinations of situations: Antecedent A: Consequent C: Conditional: "We invade Iraq" "We open the gates of hell" If A, then C (1) T T ? (2) T F F (3) F T ? (4) F F ? since we might or might not invade Iraq and might or might not open the gates of hell.

4. Reasoning with conditionals.

All of these appear to be "true possibilities". Their true values are not known at this time. Which are "those literal propositions (affirmative or negative) in the premises that are true"? I haven't a clue. Maybe JL and B know something I don't know, but the situation that we confront here seems to be typical of ordinary reasoning, where those who support war in Iraq, as much as those who oppose it, would agree that there are indeed these four possible outcomes. Ordinary reasoning, which is hypothetical, just takes for granted--simply assumes for the sake of argument--the truth of the premises as given! Even the combination true antecedent/false consequent is a true possibility, namely: the true possibility that would show the conditional premise is actually false. But not all of these "true possibilities" have to qualify as equally relevant to specific forms of reasoning, which depend upon the precise interpretation of the conditional. The point of hypothetical reasoning is to drawn out the consequences about what would be the case, if the situation the premises describe should come to pass. I would concede that JL and B may be attempting to articulate something they are not actually saying here, namely: that ordinary reasoning involves taking for granted the truth of the premises. But of course that does not make the premises true, no more than that a person z takes an argument to be valid makes the argument valid. There remains the possibility they think that the antecedents and the consequents of the conditionals that occur in ordinary reasoning are

always related by connections of meaning ("If he is a bachelor, then he is unmarried"), of causation ("If you strike the match, then that will bring about its lighting"), or of law ("If it's copper, then it's a good conductor").

5. The subjunctive conditional.

Such an attitude might justify the position that only some, but not all, of the rows of this truth table are relevant or matter to the truth of the conditional. That would make sense if the conditional were asserted to be a subjunctive, for example, about what would happen if we were to invade Iraq. Subjunctive are justified on the basis of considerations of meaning, causation, and natural law. JL and B preclude this response by declaring the conditional of ordinary discourse to be the material conditional (forthcoming). While the subjunctive interpretation could be used to maintain only rows where the antecedent is true--(1) and (2)--actually matter, the material interpretation dictates that all four rows--(1) thorough (4)--are relevant, defeating that defense. JL and B appear to have adopted an inconsistent position, namely: the conjunction of their principle of truth with the conditional as material. Indeed, if we assume arguments that are the same except for incorporating different but logically equivalent premises, then since sentences of forms "p --> q", "-q --> -p", and "-p v q" are logically equivalent, they warrant alternative arguments that yield the same conclusion. Given an additional premise of the form "p", they support an inference to a conclusion of the form "q" by modus ponens, by modus tollens, and by negative elimination, respectively. (Depending upon the rules that govern transitions from one line to another, counterpart formal proofs require one step by modus ponens and more by modus tollens or negative elimination.) While it seems plausible that such reasoning would be relatively simple for most subjects, JL and B have recorded considerable disparities in performance, which are difficult to explain if ordinary conditionals really are material, but perhaps not if they are not.

6. Syntax and semantics.

The principles and patterns studied by logicians are typically differentiated into subcategories of syntactical and semantical rules, where syntactical rules are formal in kind insofar as they apply to arguments on the basis of their forms. This presupposes a methodology for translating arguments from ordinary language into a standardized form that permits their evaluation on the basis of syntactical rules. The justification of these syntactical rules themselves, however, is semantical, where considerations of meaning and of truth and falsity provide the foundation for their acceptability. Ambiguity, for example, must be prohibited to preclude validating inferences from "Harry is mad" to "Harry is angry" or to "Harry is insane", indifferently. This condition is known as the Requirement of a Univocal Interpretation and underlies the use of formal rules in logic. Among the most elementary considerations that apply in sentential logic or SL thus turns out to be the principle that an argument is syntactically valid in SL if and only if its corresponding conditional is a tautology. A tautology, of course, is a sentence that cannot possibly be false by virtue of its logical form, assuming a uniform interpretation. Sentences of the forms, "p v -p" (read "p or not-p"), "p --> p" (read "if p then p"), and "-(p & -p)" (read "not both p and not-p") are elementary illustrations. Their truth tables

allow of no set of assignments where sentences having these forms can possibly be false. Thus, an argument whose translation satisfies the logical form, "'p --> q' and 'p'", which allows "q" to be detached and asserted separately, will be valid just in case modus ponens is syntactically valid in SL, which will be true just in case its corresponding conditional, i.e., "[(p --> q) & p] --> q", is a tautology and cannot possibly be false, which is the case.

8. Models vs. rules.

Similar considerations apply for modus tollens, affirmative disjunction, negative disjunction, and other forms of argument studied by JL and B. Their primary thesis is that mental models are semantic while formal rules are syntactic, where semantics takes precedence over syntax in ordinary reasoning by subjects untutored in logic. They seem to believe that their "models" are not "rules", yet it should be apparent that the condition that an argument is possible-for-z (probable-for-z, valid-for-z) just in case its conclusion is true in some (most, every) model in which its premises are true is a rule! (Indeed, it is a set of rules!) It may be that JL-B wanted to separate "calculation" from "deduction" precisely because it is such an obvious example of habitual rule following in reasoning. Acknowledging as much, however, might detract from the plausibility of their position. Although the difference is not always clearly understood, persons can act (think, reason) in accordance with a rule even when they are not basing their performance on that rule. Most subjects "untutored in logic", for example, would be blissfully unaware that, in reasoning from "If we invade Iraq, then that will open the gates of hell in the Middle East" and "We invade Iraq" to "That will open the gates of hell in the Middle East", they are reasoning in accord with modus ponens. Moreover, precisely how the degree of difficulty of specific inferences should be properly measured is not apparent from their work. When arguments are formalized in systems of natural deduction, for example, rather than axiomatic systems or tree systems, the number of steps required can differ considerably (Fetzer 1993, p. 347). These considerations, taken seriously, substantially undermine the claims for the performative superiority of mental models.

9. Subjective deviations.

If reasoning in accord with a rule counts as "rule-based reasoning", then there may be far less substance to the alleged differences between mental models and rule- based reasoning. No doubt, the model of rule-based reasoning that JL and B have in their own minds corresponds to the classroom expectations of instructors in courses in symbolic logic, where the validation of proofs takes place by demonstrating that every line is either given or obtained from preceding lines using the rules where the last line is the desired conclusion. Q.E.D. Since learning proof procedures is very difficult for most students, this alone suggests that we do not reason by constructing proofs. But it leaves open the possibility that we reason in accordance with similar rules when our reasoning is proper or valid, which is capable of formal demonstration. Indeed, as the study of the practices exemplified by members of some specific community of subjects, the psychology of reasoning focuses especially upon the extent to which the principles of inference and the patterns of reasoning they display satisfy or fail to satisfy the normative standards of logic. Here the evidence is rather disconcerting. Johnson-Laird and Byrne (1991), for

example, have reported that their research subjects displayed a success rate of 91% with modus ponens, of 64% with modus tollens, and of 48% for affirmative disjunction, and only 30% with negative disjunction. Their failure rates were thus 9% with modus ponens, 36% with modus tollens, 52% with affirmative disjunction, and 70% with negative disjunction. JL and B have even recommended the exclusion of modus tollens from current psychological theories of how the mind works, because it is "too difficult" (1999, p. 112). Such an approach also dictates excluding affirmative disjunction and negative disjunction, whose success rates are even worse.

10. Is modus tollens that difficult?

This attitude--that principles of reasoning that are frequently violated should be dropped from current psychological theories of how the mind operates--strikes me as peculiar and even inconsistent with previous research, some of it by Johnson-Laird himself. Peter Wason and Johnson-Laird (1972), for example, reported that even six- and eight-year-old children have displayed a remarkable capacity for valid deductive reasoning. Thus, a child is asked, "If that boy is John's brother, then he is ten years old. He is not ten years old. Is he John's brother?" The responses of their subjects were usually correct. This problem, of course, involves a modus tollens inference. If this is a problem that six- and eight-year-old children can handle with ease, then why are JL and B reporting that it is "too difficult" for current psychological theory? Indeed, one begins to wonder if reasoning generally is simply too difficult to be included in current psychological theories of how the mind operates. That might help to explain the difficulties that Americans have displayed in understanding some crucial inferences of our time: "If Gore won the election, then he is President. But Gore is not President. Did he win the election?" (modus tollens); "The Supreme Court wrongly intervened in Florida or Gore now occupies the White House, but not both. The Supreme Court wrongly intervened in Florida. What follows?" (affirmative disjunction); or, "Either polls are mistaken or Gore actually won Florida, but not both. Polls are highly reliable. What follows?" (negative disjunction). If these arguments really are too difficult for ordinary reasoning, perhaps this sheds needed light on how we have come to have a President of the United States a few cans short of a six-pack.

11. Alternative positions.

An alternative explanation, which Over and Evans (forthcoming) adopt, is that the ordinary conditional is not the material conditional. They observe that reasoning in most contexts does not concern premises that are known to be true or believed to be true, but rather situations or states of affairs that might or might not occur, as in the case of war in Iraq. Indeed, they report, "there is good reason to think that people sometimes view premises as uncertain to some degree, even in experiments in which they are explicitly asked to make assumptions, and it takes special effort to persuade them to treat (given) premises as assumptions" (Evans and Over 1996). But if the ordinary conditional is not the material conditional but something like the subjunctive, for example, and if JL and B's "principle of truth" cannot be sustained, then it should be unsurprising if the results of their research are suspect since, ironically, they would be based on false assumptions. In

their most recent work, JL and B (forthcoming) deny the ordinary conditional is truth functional. However, as Over and Evans (forthcoming) observe, they continue to defend the so-called "paradoxes of material implication", according to which, if "p" is false, then "if p, then q" is true for any "q"; and if "q" is true, then "if p, then q" is true for any p! Consequently, if "Gore was elected" is false, then "If Gore was elected, then Crest fights cavities" must be true, but equally, "If Gore was elected, then Crest promotes tooth decay". And if "Caesar was assassinated" is true, then "If 2 + 2 = 4, then Caesar was assassinated" must be true, but so too, "If 2 + 2 = 5, then Caesar was assassinated", which is very odd. Since these only follow if the conditional "p --> q" is equivalent to "-p v q", JL and B's latest position appears to be logically incoherent.

12. A different approach.

Over and Evans (forthcoming) approach the study of conditionals as a species of hypothetical reasoning, "thought that requires representations of possible states of affairs", such as forecasting, hypothesis testing, and decision making" (Evans, Over, and Handley 2001). They too employ the language of "mental models", but embrace a broader array of model types than do JL and B, whose standard models are often no more than variations on truth tables. The notion of a "mental model" is rather general, of course, and they are surely entitled to adopt it. Their approach is highly pragmatic, insofar as they assume that mental models arise within a context of objectives, aims, or goals, that reasoners typically consider only one possibility at a time, that they weight the probabilities of alternatives with respect to their objectives, and that they practice a form of intellectual satisficing, where more complex models are preferred to simpler models only when those simpler models prove inadequate (forthcoming, pp. 10-11). According to their account, the word "if" invites listeners in communication situations to construct mental models of hypothetical situations or "worlds" in which the antecedent of the conditional is satisfied and then to consider the possibility, the probability, or the inevitability of the consequent's satisfaction within that context. They presume that, as in other standard conversational contexts, principles such as the requirement of relevance apply, where the content of conditionals introduced in discourse tends to bear upon the participants' objectives. Significantly, they adopt (what is known as) the Ramsey test for appraising conditionals, where persons who disagree about if p then q are disputing the degrees of belief that should be accorded to "q" when they accept "p" and debating the probability of q, on the assumption, p.

13. The Ramsey Test.

According to F. P. Ramsey (1931), disagreements over if p then q conditionals, such as, "If we invade Iraq, then that will open the gates of hell in the Middle East", are converting a debate over a conditional into one about the degree of belief that should be granted the consequent, q, "by adding p hypothetically to their stock of knowledge and arguing on that basis about q" (Ramsey 1931, p. 247). If degrees of belief are taken to be subjective probabilities that are suitably formalized as "P(…/__)" for the probability of "…" given "___", then they are debating the probability of q given p or the proper value of P(q/p). Subjective degrees of belief do not have to satisfy the axioms of probability,

for which P(q/p) + P(-q/p) = 1, since "q" and "-q" are mutually exclusive and exhaustive, for example, nor the principle that exclusive outcomes are additive, such that, if P(q1/p) = r1 and P(q2/p) = r2, then P(q1 v q2/p) = r1 + r2. When they do satisfy the axioms of probability, however, they are called "personal" as opposed to "subjective" probabilities. Because of the multiplicity of interpretations of probability we shall consider, it may be useful to offer a more adequate mode of formalization of personal probabilities. Since they are relative to and vary with different persons, one person z at two different times t1 and t2 and two different persons z1 and z2 at the same time t may ascribe very different values to q, given p. Making the relativity to persons and times explicit, the notation, "Pzt(q/p) = r", stands for "the personal probability P for person z at time t of q, given p, equals r", which is taken to be equivalent to "Pzt.p(q) = r" as "the personal probability P for z at t of q when z accepts p, equals r"; or ,alternatively, "[P(q/p) = r]zt" and "[Pp(q) = r]zt", in order to identify them as "personal" probabilities (Fetzer 1983a).

14. Personal probabilities.

It should be observed that the very notion of "personal probabilities" carries a significant normative dimension, since there are no laws of nature, genetic or cultural, that require persons distribute their "degrees of belief" in accordance with the axioms of probability. Indeed, attributing the existence of personal probabilities implies that persons satisfy a condition known as "coherence", which entails that their subjective probabilities are mathematically tractable in these and other ways (Fetzer 1981, pp. 274-287). The ordinary axioms of probability identify conditional probabilities of the form, P(q/p), with others, such as P(q&p/p), which entails a standard mathematics for which P(q/p) = P(p/q) x P(q)/P(p), known as "Bayes' theorem" (Fetzer 1981, pp. 204- 208). The mathematics becomes complicated very fast, which suggests that there may be vastly fewer instances of "personal probabilities" than there are of degrees of belief. One of the benefits from adopting the Ramsey test as an avenue of approach toward these matters, however, is that it yields explanations for puzzling phenomena that have been uncovered by the Wason selection task (Wason 1966). This research supports the finding that subjects are far less competent in dealing with conditionals when those concern abstractions such as numbers or letters that are remote from daily experience. They also reflect the tendency to dismiss cases in which antecedents, such as "We do not invade Iraq", as irrelevant when evaluating conditionals such as, "If we invade Iraq, then that will open the gates of hell in the Middle East", which confirms Over and Evan's conclusion that ordinary conditionals are not material (forthcoming). It supports the inference that these conditionals are hypothetical or subjunctive instead.

15. Hypothetical reasoning.

Applying the Ramsey test to conditionals, persons accept their antecedents (such as "We invade Iraq") and then debate the probability of their consequents (such as "That will open the gates of hell in the Middle East"). Clearly, one person at two times and two persons at one time can accept an antecedent hypothetically--for the sake of argument--without agreeing on the probability of the consequent, especially in light of access to new data, information, or evidence, when its truth or falsity (or presence or absence) makes a

difference to those probabilities. One person might maintain that taking decisive military action against Saddam Hussein will rally other Arab states to our cause. Another might maintain that precisely the opposite effect is to be expected, where Arab states are surely going to assist another Arab state in resisting Western imperialism. They are debating the probable effects of taking the action described by the antecedent, given their other beliefs. As Over and Evans also observe, there are various kinds of data, information, and evidence that can affect subjective assessment of the respective probabilities of outcomes. One might be information about relative frequencies with which instances of the kind that is described by the antecedent have been correlated with effects of the kind described by the consequent. Another might be memories or recollections of other events in which an outcome of one kind occurred rather than another, where these memories or recollections influence subjective assessments because, unlike frequency data, they are easily available (Kahneman, Slovic, and Tversky 1982). There is obvious potential for conflict, insofar as available memories may differ from frequency data, where those recollections may be too small in number or too peculiar in circumstance to support drawing generalizations. Different background beliefs and habits of mind may preclude settling any disagreement.

16. The Principal Principle.

The distinction between anecdotal evidence and clinical trials may be sufficiently well-known that most educated persons would grant greater evidential weight to results from systematic clinical trials over unreliable experiential reports. Available evidence might be anecdotal and unreliable, while more reliable clinical data may be unavailable. The question thus arises of what kind of evidence ought to be taken into account in the assessment of subjective probabilities. Granting that anything that makes a difference to subjective probabilities qualifies as "relevant" in some broad sense, what kinds of data, information, or evidence should make a difference to the formulation of these subjective probabilities? Alternatively, if we assume that these subjective probabilities are personal, then what kinds of objective evidence should be a foundation for establishing their values? An important answer to this question has been proposed by David Lewis (1983), who advances (what he calls) "the Principal Principle" to resolve it. According to Lewis, when the objective chance for an outcome is known--when, let us say, the chance for an outcome B, under conditions of kind A, belongs to z's background knowledge--then the appropriate value for z's personal probability that B will occur, given A, should equal the objective chance for B to occur, given A. Everything else z knows qualifies as irrelevant. If that objective chance is represented by "P(B/A) = r", then "[PP(B/A) = r & A(B) = r]zt", i.e., the subjective probability of B, given A, for z at t when z knows the objective probability for B, given A, equals r, is r. So z's personal probabilities should equal objective chances when they are known. And this, says Lewis, is all we need to know about objective chance. But, as will become apparent, even he has been forced to concede that this claim is not true.

17. Another representation.

This sounds plausible and even implies a standard representation for inferences of this kind by means of a familiar argument schema for probabilistic arguments that take a

statistical generalization or a probabilistic law as a "covering law" to subsume a specific case. If we continued to use "P(B/A) = r" for objective chance, then that schema is this:
(S-1) P(B/A) = r
A
========== [r]
B
where the double-line between premises and conclusion indicates that this is an inductive argument whose conclusion can be false even when its premises are true. Schemata of this kind presume that the value of "[r]" should equal the value of "r" when that value is known and that the description of "A" includes all of the properties whose presence or absence makes a difference to the outcome described by "B" (Fetzer 1981, esp. Part II). When the value of "[r]" is taken to be the same as the value of "r", "[r]" qualifies as a logical probability. The identification of values could be justified on the grounds that the relative frequency for the conclusion to be true when the premises are true is equal to the relative frequency for an outcome of that kind to occur under conditions of that kind as a matter of objective chance. However, the value of "[r]" does not have to correspond to the value of "r". Schema (S-1) represents the relationships that ought to obtain between objective chances and personal probabilities assuming that its premises belong to a person z's beliefs and that z acts in accord with Lewis' Principal Principle.

18. Subjective deviations.

Suppose, for example, that the objective chance of obtaining heads as the result of the next toss with this (homogenous and symmetrical) coin equals _. Then Schema 1 could be implemented by means of a corresponding scheme for that specific coin toss; (S-2)
P(heads/toss with coin) = _
This is the next toss.
========== [_]
Its outcome is heads.
So far, so good. But suppose instead the situation concerned the mother of a child who was about to undergo an operation for which the objective chance of survival equals .7. In that case, of course, by the axioms of probability, the objective chance for death is .3. Yet, it is not difficult to imagine that, even though her child knew the objective chance for survival equals .7, he could not even bear the thought his mother might not survive. Under these circumstances, the value z would assign to survival as the outcome of his mom's operation might deviate arbitrarily from its objective chance up to 1, say: (S-3)
P(survival/operation) = .7
This is my mom's operation.
========== [.99]
My mom survives.
where the difference between the logical probability that should be assigned and the personal probability actually assigned represents a measure of the influence of non-logical factors. When the values of personal probabilities deviate from the values of corresponding logical probabilities, they represent degrees of "irrationality of belief".

19. Rationality of belief.

Broadly speaking, rationality of belief consists in the tendency to proportion your beliefs in accordance with the available relevant evidence, accepting hypotheses that are confirmed, rejecting hypotheses that are disconfirmed, and leaving in suspense--neither accepting nor rejecting--hypotheses that are neither confirmed nor disconfirmed. This conception takes for granted that, with respect to confirmation and disconfimation, the available evidence is sufficient to justify acceptance or rejection, which depends on the principles of reasoning that one ought to accept. The alternatives range across several conceptions of science, including Inductivism, Deductivism, Hypothetico-Deductivism, and Abductivism (Fetzer 2002b). Different sets of hypotheses may be acceptable or rejectable as a function of the standard of rationality that determines these outcomes. The role of the Principal Principle, accordingly, is to impose constraints upon the beliefs or degrees of belief that are appropriate within the context of rationality of belief. The differences in value between logical probabilities that ought to be assigned and personal or subjective probabilities that actually are assigned to various outcomes by specific persons zi at times tj thus constitute measures of the extent to which zi at tj manifests irrationality of belief. There are ample opportunities here for psychological research. In the case of schema (S-3), for example, z's degree of irrationality would be equal to the difference between the logical probability of .7 and the personal probability of .99, which is a measure of .29. No doubt, tendencies toward irrationality are strongly affected by lack of education, level of intelligence, and degree of emotional involvement. But, without an appropriate normative standard, these deviations could not be measured.

20. Objective chances.

It comes as some surprise, therefore, that the very idea of objective chance is fraught with ambiguity. There are at least three candidates for that role, including the classic conception elaborated by Laplace, according to which the probability P of an outcome B is equal to the number of favorable outcomes, B, relative to the number of equally possible outcomes, A. This approach exerts wide appeal for games of chance, including those with dice and cards, since it appears to produce the right results: the probability for drawing an ace with a standard deck given a random shuffle turns out to equal the number of aces in the deck (4) divided by the number of cards (52) which are equally possible, which is equal to 1/13. But when the dice are loaded or the deck is stacked, the results provided by the classic conception do not appear to be correct. A more appealing approach has been the frequency conception advanced by Richard von Mises and Hans Reichenbach, according to which the probability P of an outcome, B, within a reference class, A, is equal to the relative frequency which that outcome occurs in that class. Provided that every property F that makes a difference to outcomes of kind B under conditions of kind A has been taken into account, then even when the dice are loaded or the deck is stacked, the frequency conception still applies. Indeed, it can even be extended for infinite sequences by identifying probabilities as the limiting frequencies with which outcomes would occur in those sequences, were they extended indefinitely across time. While relative frequencies can exist when a reference class is finite, limiting frequencies exist only when such a class is infinite, where its members have to be ordered as a sequence (Fetzer 1981, 2002a, and 2002b).

21. Lewis' Humean epistemology.

A problem arises at this juncture for Lewis' Principal Principle, at least when it is combined with his Humean view of epistemology. According to David Hume, the only properties and relations in which we are rationally entitled to believe are those thar arise in impressions from direct experience. Our conception of causation, he maintains, brings together three ingredients, namely: instances of the kind cause (strikings of matches, say) and instances of the kind effect (lightings of those matches) resemble one another; they occur in spatio/temporal proximity to one another (with strikings preceding lightings); and there is an underlying necessary connection by virtue of which those strikings bring about those lightings. Hume contended that, while resemblance relations and relations of spatio/temporal proximity are accessible to experience, necessary connections are not. Consequently, we must abandon the notion of necessary connection and reduce the idea of causation to relations of resemblance and of spatio/temporal proximity, exclusively. Hume admitted that we do anticipate the future on the basis of regularities that have obtained in the past, but that there is no rational basis for such a habit of the mind. This means that, on a Humean analysis, laws report regular associations between events, rather than natural necessities. Consequently, deterministic causal laws are reducible to constant conjunctions and indeterministic causal laws to relative frequencies. Because he wants to based his Principal Principle on objective chances and denies that information of such a kind can include the future history of the world--which, on Hume's principles, we cannot possibly know--Lewis must revise his principle to avoid taking for granted what we cannot possibly know, namely: the future course of history, including, especially, the limiting frequencies with which events occur during the world's history (Lewis 1994).

22. The propensity conception.

As it happens, there is an alternative that, while violating Hume's constraints on the rationality of belief in necessary connections, provides an approach that salvages the Principal Principle. According to the propensity conception elaborated by Karl Popper, which I have refined, the probability P of an outcome B under conditions A is equal to the strength of the causal tendency for events of kind A to bring about outcomes of kind B (Fetzer 1981, 1988, 2002b). Propensities are single-case probabilities that bring about one or another outcome within a fixed class of outcomes under constant conditions. They produce frequencies when those conditions are subject to replication, which are evidence for inferences about their values and which, in turn, they can explain. Finite short runs and infinite long runs are merely finite and infinite sequences of repeatable singular trials. Lewis himself seeks to preserve the tenability of his principle by adopting (what he takes to be) a Humean variation, according to which propensities are properties of the world W at time t. So long as the kind of information required by the Principal Principle is current rather than future, Lewis believes that he can reconcile knowledge of objective chances with his Humean epistemology (Lewis 1994). But that is an impossible task on several grounds. First, if propensities were properties of worlds at times, as he maintains, then they would be unique and irrepeatable (Fetzer 1983b). Second, because propensities are subjunctively conditional (concerning the strength with which something would happen if

specific conditions were replicated), they represent non-Humean necessary connections. Third, it is because they are repeatable and represent non-Humean necessary connections that they possess explanatory and predictive significance (Fetzer 1981, 1983, 1988, 2002b). Lewis' principle is salvageable only by distinguishing different kinds of objective chance.

23. Formalizing propensities.

In order to present a syntax that differentiates between objective chances of these different kinds, we should retain the familiar formalization, "P(B/A) = p", for frequencies under either a relative or limiting frequency interpretation and introduce a new formalism, namely: "(x)(t)(Axt =r=> Bxt'", read, "For all x and all t, A-ing x at t would bring about (with propensity r) B-ing x at t'". The A conditions can be separated, as in the case of a toss with a die and tossing device, where the outcome of an ace (deuce, trey, and so forth) comes about with a fixed propensity (of 1/6 apiece, when the die is not loaded). Loading a die alters its propensities, even if the die and tossing device were destroyed in a fire and never tossed once! That propensity would have been responsible for the outcomes which would have been produced, if that die and tossing device had been subjected to repeated trials, even though such events never occurred. And similarly for like dice and devices. So the formula presented by Schema (S-1) needs to be replaced by a counterpart for propensities in lieu of frequencies in order to preserve the tenability of the principle: (S-4) (x)(t)(Axt =r=> Bxt')

Abt1
========== [r]
Bbt1'

where "Abt1" describes a specific instance of "A" instantiated by b at t1 and "Bbt1'" a specific instance of "B" instantiated by b at t1', where t1' = t1 + delta t (a fixed interval). The considerations that warrant the identification of the value of "r" with the value of "[r]" are now stronger than before, since "r" represents the strength of the tendency for an outcome of kind B to occur on a single trial of kind A rather than a sequence of trials.

24. Propensities vs. frequencies.

Some examples that illustrate their difference may help to bring home precisely why the distinction between frequencies and propensities matters. Consider. A 76-year- old woman was killed near Duluth 15 July 1993 when a tree that had been gnawed by a beaver fell and landed on her while she canoed the Bruile River. "Emergency personnel were called to a boating accident, but it was 'a freak of nature' accident instead, Sheriff Marvin Arneson said. The tree apparently had been gnawed by the beaver the previous night. It fell and hit the victim on the head as she and her daughter paddled past, he said. The woman was pronounced dead at the scene" (Taylor 1993). Canoeing on a river does not ordinarily result in death, but under these precise conditions, it would. The improbability of death from canoeing as a frequency is not an appropriate guide. Sometimes unusual conditions produce a close call rather than death. A 21- year-old man was gunned down by a pistol-wielding mugger, but a metal cigarette lighter he was carrying in his front pants pocket deflected the .25-caliber bullet. "The Greek exchange

student was waiting at a Chicago train station when he was accosted by three thugs who demanded his leather jacket. (Nikolas) Patronis, who doesn't speak English well, didn't understand them . . . so the impatient gunman shot him" (National Enquirer 1993). Others are not so lucky. A Dekalb, NY, man who walked away after his truck hit a utility pole was electrocuted moments later as he stepped on two downed power lines. "A pickup driven by Bradley J. Hodgdon, 23, ran off the road Saturday evening and smashed into a utility pole, shearing the pole in half, ... Hodgdon apparently survived the accident without injury, but was killed as he walked toward the road when he stepped on two 4,800-volt power lines" (News Tribune 1993).

25. Ordinary conditionals.

The conception of the ordinary conditional as the material conditional, I submit, cannot be sustained. A family of subjunctive conditionals prevail in ordinary discourse, from those based upon relations of meaning ("If he were a bachelor, then he would be unmarried"), of causation ("If you were to strike that match, then that would (probably) bring about its lighting"), and of law ("If it were copper, it would be a good conductor"). Disagreements about conditionals based upon relations of meaning are usually resolved by resorting to the dictionary, though special cases--like the meaning of "probability"--may require more extensive research. Debates about laws are often solved by consulting experts or by conducting experiments to confirm them. But arguments about causation in history commonly revolve about the specific conditions present in a particular case, such as the prospect of invading Iraq. The expected consequences are not completely known. The truth of a conditional, moreover, depends upon the complete set of conditions instantiated in the given case. If the values of relevant variables are not specified or if the values specified are not correct, then the subsuming conditional may be the wrong one or even a false generalization. Lewis proposed a semantics for subjunctives for which the truth of an antecedent and its consequent are sufficient conditions for the corresponding subjunctive: "if 'p & q' is true, then if p were, then q would be' is true', where "p & q" entails "p ==> q", using the double-arrow. ". . . ==> ___" as the subjunctive conditional. Such a principle, however, hopelessly obscure the difference between natural laws and mere correlations and cannot possibly be satisfied in an indeterministic universe. Thus, "Caesar is assassinated" and "Crest fights cavities", are both true. Yet it would be absurd to infer, "If Crest were to fight cavities, then Caesar would be assassinated" (Fetzer 1981).

26. Psychological research.

That psychologists have sometimes been misled by philosophers does not appear to be difficult to demonstrate. JL and B, for example, take for granted that Lewis' account of subjunctive and counterfactuals--subjunctives with historically false antecedents--can be taken for granted: "theorists generally take counterfactuals to be true if, in any situation in which the antecedent is true, and which otherwise resembles the real world as closely as possible, the consequent is also true", citing Stalnaker 1968, 1981; Lewis 1973; and Rips and Marcus 1977 (1991, p. 64). But the absurd results we have encountered indicate that even these worthies can sometimes make mistakes. Over and Evans (forthcoming) do far

better by taking ordinary reasoning as hypothetical reasoning that assumes antecedents to be true within a context of background knowledge that focuses on what must also be true (or probably will be true) on the assumption that those antecedent were satisfied, where the only other parts of the world deserving consideration are those that make a difference. Over and Evans (forthcoming) conclude that the ordinary conditional is not the material conditional, which appears indisputable. They also report that subjects are not concerned with situations in which the antecedent is not satisfied when they construct mental models of conditionals. This also appears to be correct and, I suggest, not only supports the characterization of ordinary reasoning as hypothetical but further confirms the suspicion that ordinary conditionals are subjunctive (albeit of several species based on relations of meaning, of causation, and of natural law). They endorse the Ramsey test as an appropriate model of ordinary reasoning, which also appears to be correct, as does a satisficing strategy that seems to prevail in hypothetical reasoning (forthcoming).

27. Future collaboration.

Over and Evans (forthcoming) welcome collaboration between psychologists and philosophers, which philosophers, I am sure, welcome in return. For example, they report that the philosopher William Lycan maintains that, when persons assert conditionals of the form, if p then q, they mean that q holds in ever possibility p that they "have at least tacitly in mind as a live prospect" (2001, p. 19). Lycan adds, they report, that persons who make such assertions cannot take p & -q as a real possibility. Over and Evans (forthcoming) reply that their subjects have commonly debated the truth of conditionals, where those that affirm them attribute higher probabilities to the truth of their consequents, given their antecedents, than do those who deny them. In their view, Lycan makes a mistake in denying that p & -q is ever a real possibility; indeed, by debating the issue, these subjects are conceding that very real possibility. While Lycan (2001) may be correct about categorical assertions of truth for conditionals based upon meaning and deterministic causal relations, he cannot be correct for indeterministic causal relations or for conditionals whose truth is not yet known. There appears to be a contextual difference at stake here, where those who argue about the truth of ordinary conditionals agree that either q or -q is possible. The truth of causal and lawful conditionals depends upon the condition that all of the causally or lawfully relevant properties that make a difference must be affirmed or denied by their antecedents, while the adequacy of explanations also depends on the exclusion of irrelevant factors (Fetzer 1981), which may be the root of satisficing. The debates that typify ordinary discourse revolve about which factors are relevant and irrelevant, where the truth of the conditional hinges on precisely this question.

28. Collaborative responsibility.

Under these conditions, I would be remiss in collaborative responsibility were I not to observe that some important concepts may or may not satisfy the axioms of probability. The propensity conception, for example, affirms a causal connection between A and B, which means that these probabilities are not a kind that satisfies the standard axioms, including conditional probabilities, for which P(B/A) = P(A&B)/P(A) = P(A/B) x P(B) =

P(B/A) x P(A). These formulations take for granted that probabilities are symmetrical, where the existence of a value for p given q entails the existence of a value for q given p. But if one is cause of the other, it is ordinarily not the case that the latter is also the cause of the former --which can happen with laws of coexistence. This means that propensities are not conditional probabilities and do not satisfy standard axioms (Fetzer 1981). Additionally, there is an almost irresistible temptation to take for granted that measures of evidential support must be mathematical probabilities. There are many reasons for doubting that this is the case, however, including that there will always be infinitely many hypotheses that cluster around any specific value, where it seems very odd that their measures of evidential support should sum to 1; and that there will always be an arbitrary number of similar hypotheses to any other, where it seems very odd to suppose their values should be additive (Fetzer 1981, p. 276). I shall not make the case here, but I believe that there are arguments for non-probabilistic measures of evidential support, which provide overwhelming reasons for preferring an approach based on likelihoods instead (Fetzer 2002b)

29. Another point of view.

The study of the psychology of reasoning as a descriptive activity depends upon the philosophy of reasoning as a normative activity. The most important contributions that psychology can make include ascertaining where descriptive practice deviates from normative competence. None of us can establish when reasoning goes wrong if we do not know when reasoning goes right. In this regard, I suspect that philosophers have not done all they could to help their colleagues in psychology. The differences between even major students of these subjects make it quite apparent that nothing is cut and dry, where psychologists would do well to keep an open mind. In my judgment, JL and B have been overly influenced by standard work in logic that simply does not apply to ordinary conditionals. About the only case in which we use conditionals of that kind is in assertions like, "If I'm wrong about this, then I'm a monkey's uncle", meant as a strenuous affirmation that I'm not wrong about this. Let's hope I'm also right now. Among the most interesting general consequences that emege from all these considerations, I believe, is that ordinary reasoning does not satisfy the standards of logic because persons are complex entities affected by their motives, beliefs, ethics, abilities, and capabilities. While students of language characteristically distinguish between the intensions and the extensions of words as the conditions that have to be satisfied for a word to apply to something and as the class of entities--past, present, and future--to which they properly apply, ordinary reasoning is powerfully affected by connotations and denotations as the emotions and attitudes we associate with things of a kind and the specific things of that kind with which we happen to be familiar. As a consequence, we too often draw hasty generalizations from unrepresentative samples.

30. Cognitive geography.

As an illustration, I recommend consideration of the area of inquiry known as "cognitive geography", which subsumes the activity of constructing mental maps of our environment. You may recall the celebrated New Yorker cover where a map of the

United States is dominated by Manhattan up to the Hudson River, where Chicago and San Francisco appear somewhere in the mid- and far-west, which are compressed beyond recognition. Except that, as a representation of the emotions and attitudes of the residents of New York City, it expresses precisely how they feel about themselves and their environment. Manhattan and its distinctive features are crucially important to those who live there. Chicago and points west are relatively insignificant, because they do not exert a profound influence on getting around in life when lived in the city. Anyone who doubts that human beings are not digital machines ought to find this result intuitively satisfying. Computer models of the mind, which predominate cognitive science as a discipline, should imply that reasoning, especially deductive, is an exemplar of what minds can do best. The findings of JL and B and of Over and Evans suggest that the reality does not support that expectation. This appears to be one more prediction that undermines confidence in the prevailing paradigm (Fetzer 1990, 1996, 2001). The humbling discovery that we are like anything but machines in our mental modes of operation may appear to be a slight advance in human understanding, but it contributes to understanding what it is to be human.

6 December 2002 References

Evans, J. St. B. T. and D. E. Over (1996), Rationality and Reasoning (Hove, UK: Psychology Press).
Evans, J. St. B. T., D. E. Over, and S. H. Handley (2001), "A Theory of Hypothetical Thinking", in D. Hardman and L Maachi, eds., Decision Making and Reasoning (Chichester, UK: Wiley and Sons), pp.
Fetzer, J. H. (1981), Scientific Knowledge (Dodrdrecht, Holland: D. Reidel).
Fetzer, J. H. (1983a), "Probability and Objectivity in Deterministic and Indeterministic Situations", Synthese 57 (1983), pp. 367-386.
Fetzer, J. H. (1983b), "Probabilistic Explanations", in P. Asquith and T. Nickles, eds., PSA 1982, Vol. 2 (East Lansing, MI: Philosophy of Science Association), pp. 194-207.
Fetzer, J. H. (1988), "Probabilistic Metaphysics", in J. H. Fetzer, ed., Probability and Causality (Dordrecht, The Netherlands: D. Reidel), pp. 109-132.
Fetzer, J. H. (1991), Artificial Intelligence: Its Scope and Limits (Dordrecht, The Netherlands).
Fetzer, J. H. (1993), "The Argument for Mental Models is Unsound", Behavioral and Brain Sciences 16 (1993), pp. 347-348.
Fetzer, J. H. (1996), Philosophy and Cognitive Science, 2nd edition (St. Paul, MN: Paragon).
Fetzer, J. H. (1999a), "Deduction and Mental Models", Minds and Machines 9 (1999), pp. 105-110.
Fetzer, J. H. (1999b), "Mental Models: Reasoning without Rules", Minds and Machines 9 (1999), pp. 119-125.
Fetzer, J. H. (2001), Computers and Cognition: Why Minds are Not Machines (Dorecht, The Netherlands: Kluwer).
Fetzer, J. H. (2002a), "Introduction", Synthese 132 (2002), pp. 5-8.
Fetzer, J. H. (2002b), "Propensities and Frequencies: Inference to the Best Explanation", Synthese 132 (2002), pp. 27-61.
Johnson-Laird, P. (1983), Mental Models (Cambridge, UK: Cambridge University Press).
Johnson-Laird, P. and R. M. Byrne (1991), Deduction (Hillsdale, NJ: Earlbaum).

Johnson-Laird, P. and R. M. Byrne (1993), "Mental Models or Formal Rules?", Behavioral and Brain Sciences 16 (1993), pp. 368-380.
Johnson-Laird, P. and R. M. Byrne (1999), "Mental Models Rule, OK? A Reply to Fetzer", Minds and Machines 9 (1999), pp. 111-117.
Johnson-Laird, P. and R. M. Byrne (forthcoming), "Conditionals: A Theory of Meaning, Pragmatics, and Inference", Psychological Review (forthcoming).
Lewis, D. (1973), Counterfactuals (Oxford, UK: Basil Blackwell).
Lewis, D. (1983), "A Subjectivist's Guide to Objective Chance", in R. Jeffrey, ed., Studies in Inductive Logic and Probability, Vol. II (Berkeley, CA: University of California Press, 1983), pp. 267-297.
Lewis, D. (1994), "Humean Supervenience Debugged", Mind 103 (1994), pp. 473-490.
Lycan, W. (2001), Real Conditionals (Oxford, UK: Oxford University Press).
National Enquirer (1993), "Cigarette lighter saves man from a bullet", National Enquirer (6 July 1993), p. 21.
News Tribune (1993), "Man electrocuted after car crash", Duluth News Tribune (11 October 1993), p. 2D.
Over, D. E. and J. St. B. T. Evans (forthcoming), "The Probability of Conditionals: The Psychological Evidence", Mind & Language (forthcoming).
Ramsey, F. P. (1931), The Foundations of Mathematics and Other Logical Essays (London, UK: Routledge and Kegan Paul).
Rips, L. J. and S. L. Marcus (1977), "Suppositions and the Analysis of Conditionals", in M. A. Just and P. A. Carpenter, eds., Cognitive Processes in Comprehension (Hillsdale, NJ: Lawrence Erlbaum), pp.
Stalnaker, R. (1968), "A Theory of Conditionals", in N. Rescher, ed., Studies in Logical Theory (Oxford, UK: Basil Blackwell), pp. 98-112.
Stalnaker, R. (1981), "A Defense of Conditional Excluded Middle", in W. Harper et al., eds., Ifs: Conditionals, Belief, Decision, Chance, and Time (Dordrecht, The Netherlands: D. Reidel), pp.
Taylor, C. (1993), "Woman canoeing Brule River is killed in freak accident", Duluth News Tribune (16 July 1993), p. 1A.
Thau, M. (1994), "Undermining and Admissibility", Mind 103 (1994), pp. 491-503.
Wason, P. (1966), "Reasoning", in B. M. Foss, ed., New Horizons in Psychology I (Harmandsworth, UK: Penguin), pp. 106-137.
Wason, P. and P. Johnson-Larid (1972), The Psychology of Reasoning (Cambridge, MA: Harvard University Press).