Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making

Edited by

David Hardman
London Metropolitan University, UK
and
Laura Macchi
University of Milan-Bicocca, Italy

JOHN WILEY & SONS, LTD

CHAPTER 6

Probabilities and Pragmatics in Conditional Inference: Suppression and Order Effects

Mike Oaksford
Cardiff University, UK
and
Nick Chater
University of Warwick, UK

INTRODUCTION

Over the last few years, we have been developing a probabilistic approach to human reasoning which suggests that many of the so-called errors and biases seen in deductive reasoning are the result of applying everyday uncertain reasoning strategies to these laboratory tasks (see Chater & Oaksford, 2000, 2001). We initially applied this approach to the Wason selection task. We argued that participants are making decisions about whether the benefits of selecting certain types of data, in terms of information gain (indicative task) or utility (deontic task), outweighed the costs (Oaksford & Chater, 1994, 1995a, 1996, 1998a, 1998b; Oaksford et al., 1997; Chater & Oaksford, 1999a; Oaksford, Chater & Grainger, 1999). More recently, we have applied this approach to syllogistic reasoning (Chater & Oaksford, 1999b).

Both of these inferential forms are complex compared to the standard conditional inference task, where participants are provided with a conditional premise *if p then q* and a range of categorical premises, *p*, not-*p*, *q* and not-*q*. Inferring *q* given *p*, and not-*p* given not-*q* correspond to the logically valid inferences of modus ponens (MP) and modus tollens (MT), respectively. Inferring not-*q* given not-*p*, and *p* given *q* correspond to the logical fallacies of denying the antecedent (DA) and affirming the consequent (AC), respectively. Because it was clear that a simplified version of the probabilistic approach we applied to the selection task could also be applied to conditional inference, we have recently shown how the model can explain polarity biases (Evans, Newstead & Byrne, 1993) in the conditional inference

Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making. Edited by David Hardman and Laura Macchi.

task (Oaksford, Chater & Larkin, 2000). This bias occurs when negations are used in the antecedent and consequent of a conditional. The principal result is that people are biased towards negative conclusions. According to our model and Oaksford and Stenning's (1992) contrast set account of negations, this bias occurs because negations define high-probability categories. Oaksford et al. (2000) showed that when manipulating probabilities instead of negations, a high-probability conclusion effect is observed analogous to negative conclusion bias. Consequently, their application of a probabilistic model to conditional inference seems to explain one of the principal biases observed in a rational probabilistic framework.

However, there are other effects in conditional inference that Oaksford et al. (2000) did not address and that may be understood from a probabilistic perspective. Our goal in this chapter is to speculate on whether this is feasible. The two effects we look at are suppression effects (e.g., Byrne, 1989; Cummins et al., 1991; Cummins, 1995; Byrne, Espino & Santamaria, 1999) and order effects (Evans, 1977; Evans & Newstead, 1977; Evans & Beck, 1981; Thompson & Mann, 1995; Girotto, Mazzocco & Tasso, 1997; Evans, Handley & Buck, 1998). Suppression effects occur when further information reduces the degree to which a participant is willing to endorse an inference. For example, if you are told that *if the key is turned, the car starts* and that *the key is turned*, you are likely to endorse the MP inference to the conclusion that *the car starts*. However, if you are also told, *if the petrol tank is not empty, the car starts*, you are less likely to endorse this conclusion because the car may not start if the petrol tank is empty. The petrol tank being empty provides an *exception* to the rule. Byrne (1989), who called such cases "additional antecedents", showed that they suppress the valid inferences of MP and MT. Other information can suppress DA and AC. For example, if you are told that *if the key is turned, the car starts* and that *the key is not turned*, you might endorse the DA inference to the conclusion that *the car does not start*. However, if you are also told, *if the car is hot-wired, the car starts*, you are less likely to endorse this conclusion because the car may start because it has been hot-wired. Byrne (1989), who called such cases "alternative antecedents", showed that they suppress the fallacies of DA and AC.

Order effects occur when, for example, the order of clauses is reversed, as in the conditional *q only if p*. When this is done, participants tend to endorse more AC and MT inferences and fewer MP and DA inferences. In this chapter, we argue that our simple probabilistic model can provide quite detailed accounts of suppression effects. Our account of order effects derives the connection between conversational pragmatics and subjective probability that we first discussed in Oaksford and Chater (1995a).

However, before we turn to our account of these effects, we first outline the reasons why we believe that a probabilistic approach to conditional inference is required. This involves a brief discussion of the inadequacy of the material conditional of standard logic in providing an account of the conditional as it used in everyday inference.

LOGICISM AND UNCERTAINTY

Within philosophy, linguistics, logic and computational theory, there is general convergence on the view that standard first-order logic is inadequate to capture everyday reasoning about the real world. Although some psychologists are well aware of these literatures, we believe that their implications concerning the scope of first-order reasoning have not been fully recognised. Indeed, the very fact that the two leading formal psychological theories of reasoning, mental logic (e.g., Rips, 1994) and mental models (e.g., Johnson-Laird & Byrne, 1991), both retain the standard logical apparatus suggests that the inadequacies of first-order logic as a model for human reasoning are not universally accepted. We first sketch the standard logical treatment of the conditional, and then consider its problems and attempted solutions to these problems within a logical framework.

Problems with the Material Conditional

The standard approach within the formal semantics of natural or logical languages is to provide a recursive definition of the truth of complex expressions in terms of their parts. The natural language phrase *if p then q* is usually rendered as the material conditional of logic. The material conditional $p \rightarrow q$ is true if and only if p is false or q is true (or both). This semantics licenses the valid rules of inference, modus ponens (MP) and modus tollens (MT). There are certain well-known counterintuitive properties of this semantics. For example, it means that any conditional with a false antecedent is true—thus, the sentence, "if the moon is striped, then Mars is spotted", is true according to the material conditional. But, intuitively, it is either false or nonsensical.

Further problems arise because the material conditional allows "strengthening of the antecedent". That is, given the premise *if p then q*, we can conclude that *if (p and r) then q*, for any *r*. Strengthening of the antecedent seems appropriate in mathematical contexts. *If it is a triangle, it has three sides* does imply that *if it is a triangle and it is blue, it has three sides*. Indeed, this is a crucial feature of axiomatic systems in mathematics—axiomatisation would be impossible if adding new axioms removed conclusions that followed from the old axioms. However, strengthening of the antecedent does not apply to most natural language conditionals, which, as we have argued, are uncertain. For example, *if it's a bird, it flies* does not allow you to infer that *if it's a bird and it's an ostrich, it flies*. That is, for natural language conditionals, conclusions can be lost by adding premises; that is, strengthening the antecedent does not hold. Furthermore, note that whether some additional information *r* has this effect or not is content dependent; for example, if you learn that this bird is a parrot, the conclusion that *it can fly* is not lost. The distinction between inference systems in which strengthening of the antecedent does or does not hold is of central importance to knowledge representation in artificial intelligence. Roughly, inference systems where strengthening of the antecedent holds are known as monotonic systems (continuously adding premises leads to continuously adding conclusions, without removing any); inference systems where strengthening of the antecedent does not hold are non-monotonic. In artificial intelligence, it is universally accepted that human everyday reasoning is uncertain and thus non-monotonic, and that developing systems for non-monotonic reasoning is a major challenge (e.g., McCarthy & Hayes, 1969; Ginsberg, 1987).

Regarding our first problem with material implication, that a false antecedent guarantees the truth of a conditional, an intuitive diagnosis is that material implication fails to specify that there be any connection between the antecedent and the consequent—they can simply be any two arbitrary propositions. Within the logical literature, there have been two general approaches to capturing this intuition—relevance logic and modal logic.

Solutions?: Relevance and Modality

Relevance logic, as its name implies, demands that there be a relationship of "relevance" between antecedent and consequent, where this is defined in terms of the proof of the consequent involving the antecedent (Anderson & Belnap, 1975). From a logical point of view, systems of relevance logic are not well developed. For example, it has been very difficult to provide a semantics for relevance logics (Veltman, 1985); this means that it is not clear quite what notion of relevance is being coded by the syntactic rules used in particular relevance logics. But, in any case, the relation of relevance would not appear to be reducible to notions of proof, particularly not in everyday contexts, because the uncertain character of reasoning means that proofs are never possible. So relevance logics do not appear to be a useful direction for developing a notion of the conditional which applies to everyday reasoning. However, in the psychology of reasoning, Braine (1978) has advanced a relevance-based account, arguing that people naturally only assert conditionals when the consequent is deducible from the antecedent.

The second approach to the conditional employs modal notions, such as necessity and possibility. Syntactic systems of modal logic and so-called strict implication based on them were first suggested by C.I. Lewis (1918). Semantic theories for modal logics were developed much later by Kripke (1963), permitting an understanding of the notions of necessity and possibility that were being encoded in the syntactic rules. Specifically, Kripke provided a semantics in terms of "possible worlds". The idea is that different modal logics can be understood in terms of different relations of "accessibility" between possible worlds. In these terms, a proposition is necessary if it is true in all accessible possible worlds, and it is possible if it is true in some accessible possible worlds.

The most philosophically important account of conditionals is given by the Lewis–Stalnaker possible world semantics for the counterfactual conditional (Stalnaker, 1968; D. Lewis, 1973). A counterfactual conditional is one in which the antecedent is known to be false: for example, *if the gun had gone off, he would have been killed*. According to material implication, such claims are always true, simply because their antecedents are false. But, clearly, this cannot be correct—under most circumstances, the counterfactual *if he had stubbed his toe, he would have been killed* will be judged unequivocally false. Looking at the Lewis–Stalnaker semantics for such claims reveals all the problems that logical approaches to everyday reasoning must confront in philosophy and in artificial intelligence (AI).

The intuitive idea behind the Lewis–Stalnaker semantics for a conditional such as *if the gun had gone off, he would have been killed* is based on the idea that in the world maximally similar to the actual world but in which the gun went off, he died. Clearly, the major issue here is what counts as the world maximally similar to the actual one. One important criterion is that the physical laws are the same, so that speeding bullets still tend to kill people, the gun is pointing in the same direction, and so on—the only difference is that the gun went off in this world, whereas it did not in the actual world. But there is a vast range of specific problems with this account. For example, it is not at all clear how to construct a world where only a single fact differs from the actual world. This is problematic because for this to be true (assuming determinism) the difference in this crucial fact implies either a different causal history (the bullet was a dud, the gun was faulty, etc.) or different causal laws (pulling triggers does not make guns go off in this possible world). Moreover, a different causal history or different causal laws will have different causal consequences, aside from the single fact under consideration. Thus, it appears inevitable that the so-called maximally similar world differs in many ways, rather than just about a single fact, from the actual world. So, by changing one thing, we automatically change many things, and it is not at all clear what the inferential consequences of these changes should be. The problem of

specifying the ramifications of a single change to a world (or in an agent's knowledge about its world) is immensely difficult—in AI, this problem has been called the frame problem (Pylyshyn, 1987), and it has bedevilled AI research for the last 30 years. Hence, a theory of conditionals which presupposes a solution to the frame problem is unlikely to prove satisfactory as a basis for a psychology of conditional reasoning.

These problems aside, this semantics for the counterfactual (that is, where the antecedent—the gun going off—does not apply in the actual world) has also been applied to the indicative case (where the gun may or may not have gone off). Simplistically, the hypothetical element of an indicative statement, such as *if the gun goes off, he is dead*, seems to be captured by the same semantics—the only difference is that we do not know whether the actual world is one in which the gun goes off or not. Nonetheless, this kind of semantic account does avoid some of the absurdities of material implication. Thus, for example, sentences such as *if the moon is striped, then Mars is spotted* are now clearly false—in worlds maximally similar to the actual world in which the moon is striped, Mars will still look red. Crucially, it is intuitively clear that strengthening of the antecedent can no longer hold. For example, *if it's a bird, it flies* does not allow you to infer that if it's a bird and it's an ostrich, it flies. The worlds in which the antecedents are evaluated will clearly differ—the world most similar to the actual world in which something is a bird is not the same as the world most similar to the actual world in which something is an ostrich. In particular, in the first world, the thing will most likely fly (because most birds fly); but in the second world, the thing will not fly (because ostriches cannot fly). These examples suggest that the Lewis–Stalnaker semantics may provide a more descriptively adequate theory of conditionals than the material conditional.

However, for psychological purposes, we need an account of the formal processes that could implement this semantics. People do not have access to possible worlds—instead, they have access to representations of the world, which they can productively recombine to produce different representations of the way the world might be or might not have been. The programme of attempting to mechanise reasoning about the way the world might be has been taken up by the study of knowledge representation in AI. The starting point is the notion of a knowledge base that contains representations of a cognitive agent's beliefs about the world. This approach involves formal representations and formal proof procedures that operate over these representations which can be implemented computationally. However, it is far from clear that formal attempts in AI can adequately capture the Lewis–Stalnaker semantics.

Let us reconsider strengthening the antecedent and perhaps the best-known approach to this problem within AI. Problems for strengthening the antecedent arise when the inferences that can be made from one antecedent intuitively conflict with the inferences that can be made from another. For example, knowing that *Tweety is a sparrow* leads to the conclusion that *Tweety flies*, whereas knowing that *Tweety is one second old* leads to the conclusion that *Tweety cannot fly*. This leads to the problem of what we infer when we learn that *Tweety is a one-second-old sparrow*; that is, the problem of what we infer when the antecedent is strengthened. It is intuitively obvious that a one-second-old sparrow cannot fly; that is, that when Tweety is one second old, the possible world in which Tweety cannot fly is more similar to the actual world than any other possible world where Tweety can fly. Although this is intuitively obvious, formally, it is not obvious how to capture this conclusion. Formally, we can regard these two pieces of information as two conditional rules, *if something is a bird it can fly*, and *if something is one second old it cannot fly*. Formal proposals in AI (e.g., Reiter, 1985) appear unable to break the symmetry between these rules and specify which of these conflicting conclusions we should accept. That is, these proposals do not respect our intuitive understanding of how the Lewis–Stalnaker semantics should be applied. The point here is that in the example it is our knowledge of what the rules mean and how the world works that indicate that a one-second-old sparrow is not going to fly. More generally, it is not the formal properties of conditionals that determine the subsets of possible worlds in which they are evaluated in the Lewis–Stalnaker semantics. What matters is the content of the rules, to which the formal procedures for inference in logicist AI do not have access.

There have been various alternative proposals within the AI literature to deal with the problem of strengthening the antecedent, or default reasoning. The best known are McCarthy's (1980) circumscription, McDermott and Doyle's (1980) non-monotonic logic I, McDermott's non-monotonic logic II (1982) and Clark's predicate completion (1978). However, the problems that we have described above appear to apply equally to all of these approaches (McDermott, 1987; Shoam, 1987, 1988). Moreover, approaches based on formal logic within the psychology of reasoning, such as mental logics (e.g., Rips, 1994) and mental models (e.g., Johnson-Laird & Byrne, 1991), also fail to address these issues, because the approach they adopt formalises the conditional by the standard logic of material implication. However, as we have seen, the material conditional completely fails to capture the use of conditionals in everyday inference.

A Probabilistic Approach

We have seen that conditional inference is of fundamental importance to cognitive science, as well as to AI, logic and philosophy. We have suggested that the problems that arise in capturing conditional inference indicate a very profound problem for the study of human reasoning and the study of cognition at large. This is that much of our reasoning with conditionals is uncertain, and may be overturned by future information; that is, they are non-monotonic. But logic-based approaches to inference are typically monotonic, and hence are unable to deal with this uncertainty. Moreover, to the extent that formal logical approaches embrace non-monotonicity, they appear to be unable to cope with the fact that it is the content of rules, rather than their logical form, which appears to determine the inferences that people draw. We now argue that perhaps by encoding more of the content of people's knowledge, by probability theory, we may more adequately capture the nature of everyday human inference. This seems to make intuitive sense, because the problems that we have identified concern how uncertainty is handled in human inference, and probability theory is the calculus of uncertainty.

Before we outline how a probabilistic approach can account for some of the effects seen in experiments on conditional inference, we briefly consider how it can avoid some of the problems we have just introduced for the material conditional. From a probabilistic point of view, the natural interpretation of conditionals is in terms of conditional probability. Thus, the statement that birds fly (or more long-windedly, if something is a bird, it flies) can be regarded as claiming that the conditional probability of something flying, given that it is a bird, is high. Probability theory naturally allows non-monotonicity. If all we know about a thing is that it is a bird, the probability that it flies might be, say, .9 (P(flies|bird) = .9). However, the probability of its flying given that it is both a bird and an ostrich is 0 or nearly 0 (P(flies|bird, ostrich) = 0), and the probability of its flying given that it is both a bird and a parrot may be, say, .96 (P(flies|bird, parrot) = .96). All these statements are completely

compatible from the point of view of probability theory. So, from a probabilistic perspective, the result of strengthening the antecedent in these cases leads to intuitively acceptable results. This approach to the meaning of conditional statements has been proposed in philosophy by Adams (1966, 1975), and has also been adopted in AI (Pearl, 1988). There have been some problems raised with this probabilistic interpretation of the conditional. These concern the rather unnatural scenario in which conditionals are embedded; for example, *if (if p then q) then r* (Lewis, 1976). However, the relevance of these problems to the design of AI systems and to human cognition is unclear (Pearl, 1988). Certainly, as we now argue, they do not seem to impinge on our ability to provide probabilistic interpretations of conditional reasoning in the laboratory.

A PROBABILISTIC APPROACH TO CONDITIONAL INFERENCE

Several authors have suggested that human conditional inference has a significant probabilistic component (Chan & Chua, 1994; Anderson, 1995; Stevenson & Over, 1995; Liu, Lo & Wu, 1996; George, 1997). Here we outline Oaksford, Chater and Larkin's (2000) probabilistic computational level model (Marr, 1982) of the inferences that people should make in these experiments.

A Computational Level Model

In this model, rules are represented as 2 × 2 contingency tables, as in Table 6.1. In this table, with respect to a rule *if p then q*, $a = P(p)$, the probability of the antecedent; $b = P(q)$, the probability of the consequent; and $\varepsilon = P(\text{not-}q \mid p)$, the probability that the consequent does not occur given the antecedent. ε is the exceptions parameter, as used by Oaksford and Chater (1998b). For example, if p is *turn the key* and q is *the car starts*, ε is the probability that the car does not start even though you have turned the key. Following previous accounts (Chan & Chua, 1994; Stevenson & Over, 1995; Liu, Lo & Wu, 1996), Oaksford et al. (2000) assumed that people endorse an inference in direct proportion to the conditional probability of the conclusion given the categorical premise. The following expressions for the conditional probabilities of each inference can be derived from Table 6.1:

MP: $$P(q \mid p) = 1 - \varepsilon \quad (1)$$

DA: $$P(\neg q \mid \neg p) = \frac{1 - b - a\varepsilon}{1 - a} \quad (2)$$

AC: $$P(p \mid q) = \frac{a(1 - \varepsilon)}{b} \quad (3)$$

MT: $$P(\neg p \mid \neg q) = \frac{1 - b - a\varepsilon}{1 - b} \quad (4)$$

We show the behaviour of the model relevant to explaining suppression effects in Figure 6.1. This figure shows how the probability that each inference should be drawn varies as a function of ε, that is, the probability of exceptions. This probability is directly related to the suppression of the valid inferences MP and MT by the introduction of exceptions. If ε is high, this corresponds to a rule with many exceptions. Figure 6.1 shows that, as would be predicted, the probability that MP or MT should be endorsed falls as ε rises.

Figure 6.1 also shows that DA and AC seems to be affected by exceptions in a similar way. However, as we mentioned in the discussion, DA and AC seem to be most affected by the probability of alternatives; for example, the probability that a car can be started

Table 6.1 The contingency table for a conditional rule, *if p then q*, when there is a dependency between the *p* and *q* that may admit exceptions (ε). $a = P(p)$, $b = P(q)$, and $\varepsilon = P(\text{not-}q \mid p)$

	q	not-*q*
p	$a(1-\varepsilon)$	$a\varepsilon$
not-*p*	$b - a(1-\varepsilon)$	$(1-b) - a\varepsilon$

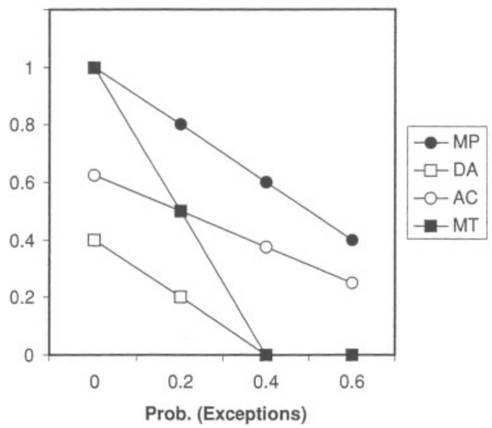

Figure 6.1 How the probability that a conclusion should be drawn varies as a function of the probability of exceptions (P[not-*q* | *p*]: Prob.[Exceptions]). The probability of the antecedent (*a* [*P*(*p*)]) and the probability of the consequent (*b* [*P*(*q*)]) were held constant at .5 and .8, respectively

other than by turning the key. This corresponds to the probability of the car's starting even though the key has not been turned; that is, $P(q \mid \text{not-}p)$. We call this probability DA′ because it is the complement of DA ($P(\text{not-}q \mid \text{not-}p)$; consequently, as the probability of DA′ increases, the probability of drawing DA decreases. DA′ is linearly related to ε, with $\frac{b-a}{1-a}$ as the intercept and $\frac{a}{1-a}$ as the slope. This reveals that, according to the model, if the probability of exceptions is kept constant, changes in DA′ will involve changes in the probabilities of the antecedent and consequent, that is, in *a* and *b* (*P*[*p*] and *P*[*q*]). We could have treated DA′ as a primitive parameter of the model. However, we chose only to parameterise exceptions for two reasons. First, we wished to be parsimonious: the model already contains three parameters. Second, linguistically, the structure of if ... then rules reflects the causal ordering of events in the world (Comrie, 1986), allowing us to predict what will happen next. These predictions go awry only because of exceptions. Thus, MP and the reasons why it might fail are particularly cognitively salient (Cummins et al., 1991), and that is why we treat exceptions as a primitive. As we will see when we model the data on the suppression effect, this choice makes experimentally testable predictions.

Oaksford et al.'s (2000) model is defined at Marr's computational level; that is, it outlines the computational problem people are attempting to solve when they are given conditional inference tasks to perform. It also specifies the knowledge that they bring to bear on the problem, that is, knowledge of the probabilities of exceptions and of the antecedents and consequents of the rules they are given. That is, we abandon the conventional view that the problem people confront is one of which logical rules to apply to these conditional statements.

EXPLAINING SUPPRESSION EFFECTS

Oaksford et al. (2000) concentrated on showing how their probabilistic model could account for polarity biases in conditional inference. However, the model can also be applied to the standard pattern of results on the task and to suppression effects. We now look at how the model applies to these data.

The Standard Results

We first show how this account explains the standard abstract data on conditional inference. To model the data, we need to find appropriate values for $P(p)$, $P(q)$ and ε. However, for the abstract material normally used in these tasks, it is difficult to know what values people would normally assume. Some constraints can be derived by looking at the base rates of natural language predicates, as it seems reasonable to assume that, when confronted with incomplete information, people will use their prior knowledge to fill in the gaps. That is, they will assume that new cases will be pretty much like those they have already seen. Most natural language predicates cut up the world into relatively small subsets; for example, most things are not birds, are not black, are not coffee pots and so on. That is, the probability that any randomly selected object is a bird, is black or is a coffee pot is very low. We exploited this "rarity" assumption in modelling Wason's selection task (Oaksford & Chater, 1994, 1996, 1998b).

However, as Oaksford et al. (2000) observe, in the context of conditional inference, a rarity assumption is probably not appropriate. They pointed out that inferences are specific to contexts. So, for example, you are only likely to need to make inferences about donkeys being stubborn in contexts where you are likely to encounter donkeys. That is, even though the base rate of donkeys is likely to be very low, in contexts where it would be appropriate to draw inferences about them, the probability of encountering a donkey is likely to be higher, possibly a lot higher than the base rate. Consequently, to model the standard data, we averaged conditional probabilities over the whole parameter space but excluding values of $P(p)$ and $P(q)$ that were less than .1, that is, the rare values. Moreover, to be able to infer reliably that a particular animal is stubborn, given that it is a donkey, there had better be more stubborn things than donkeys. If this were not the case, our inferences could go awry; that is, we would quite often encounter non-stubborn donkeys. So, for the rule *if something is a donkey* (*p*), *it is stubborn* (*q*), whatever the absolute values of $P(p)$, it is likely that $P(q) > P(p)$. To model the standard data, we therefore calculated conditional probabilities averaged only over the region of the parameter space where $P(q) > P(p)$.

To compute the predicted values, we sampled the parameter space where $P(q) > P(p)$ between .1 and .9, in steps of .1 with $\varepsilon = .1$. For each set of parameter values, we computed the predicted probability of each inference, using Equations 1–4. Overall, this meant that 36 values were calculated for each inference. Finally, we averaged over all 36 values for each inference to obtain the predicted proportion of those inferences people should make. For MP, the mean, in percentage (data mean in brackets), was 90.00 (97.36) because this probability relies only on ε. For DA, the mean was 42.86 (42.09); for AC, the mean was 45.00 (42.64); and for MT, the mean was 82.14 (62.18). Although we made no attempt to optimise these values, they agree reasonably well with the standard data, especially for DA and AC. Thus, it would appear that the standard pattern of results on the conditional inference task, which appears irrational from a logical point of view, may be a reflection of a rational probabilistic strategy. Only two assumptions were needed, (i) that participants sample quite broadly in the parameter space outside the rarity region, and (ii) that they assume $P(q) > P(p)$.

Suppression Effects

In this section, we show how our simple probabilistic model can account for suppression effects. These effects involve the way additional information, either explicitly given in the experimental set-up (e.g., Byrne, 1989) or implicitly available from prior knowledge (e.g., Cummins et al., 1991), can affect the inferential process. In either case, we assume that this information has the effect of altering the subjective probabilities that are then entered into Equations 1–4 to calculate the probability that an inference will succeed. Additional antecedents (or exceptions), for example, the information that there is petrol in the tank with respect to the rule *if you turn the key, the car starts*, concern the probability of the car's not starting even though the key has been turned—that is, they concern ε. If you do not know that there is petrol in the tank, you cannot unequivocally infer that the car will start (MP). Moreover, bringing to mind other additional factors that need to be in place to infer that the car starts—for example the battery must be charged—will increase this probability. ε is a primitive parameter of Oaksford et al.'s (2000) model and can be derived directly from the data for the MP inference. It is therefore an immediate consequence of our model that if there are many additional antecedents, that is, ε is high, the probability that the MP inference should be drawn will be low. That is, suppression of MP by additional antecedents is a direct prediction of the model.

Alternative antecedents, such as the information that the car can also be started by hot-wiring with respect to the rule *if you turn the key, the car starts*, concern the probability of the car starting even though the key has not been turned; that is, $P(q \mid \text{not-}p)$. If you know that a car can be started by other means, you cannot unequivocally infer that the car will not start although the key has not been turned (DA). Moreover, bringing to mind other alternative ways of starting cars, such as bump-starting, will increase this probability. $P(q \mid \text{not-}p)$ is DA′, the converse of DA ($P[\text{not-}q \mid \text{not-}p]$). It is therefore an immediate consequence of our model that if there are many alternative antecedents, that is, $P(q \mid \text{not-}p)$ is high, the probability that the DA inference should be drawn will be low. That is, suppression of DA by alternative antecedents is a direct prediction of the model.

Suppression effects are just what would be predicted if people regard conditional reasoning as non-monotonic and defeasible, as we discussed in the section *Logicism and Uncertainty*. There, following suggestions especially in the AI knowledge representation literature (e.g., Pearl, 1988), we proposed that conditional sentences were represented in terms of conditional probabilities. This suggestion has the consequence that, for example, the probability that *the car starts* if *the key is turned* is less than 1, say, .9; that is, P(*starts* | *key-turned*) = .9. If all we know is that *the key is turned*, the probability that *the car starts*, that is, the probability of endorsing MP, would be .9. However, the probability of its starting given that the key is turned and the petrol tank is empty is 0 (P[*starts* | *key turned, empty*] = 0). Our theoretical model amounts to applying this resolution of the conceptual problems, with strengthening the antecedent, to the empirical evidence on conditional reasoning performance.

Modelling the Suppression Effect I: Byrne (1989)

Figure 6.2 shows the overall fit of the model to Byrne's (1989) data. The observed proportions of inferences endorsed in Byrne's Experiment 1 are shown with filled diamonds (data), and the proportions predicted by the model are shown with unfilled squares (model). The three graphs correspond to the three conditions in Byrne's (1989) experiment. The *simple* condition that provided the baseline in which there was no manipulation of alternative or additional antecedents. In the *alternative antecedent condition*, participants were provided with an alternative rule; for example, *if hot wired, the car starts*. This graph reveals that this manipulation suppressed both DA and AC, but not MP or MT. In the *additional antecedent condition*, participants were provided with an additional rule; for example, *if there is petrol in the tank, the car starts*. This graph reveals that this manipulation suppressed both MP and MT, but not DA or AC. We fitted the model to the data by looking for the values of $P(p)$ and $P(q)$ that maximised the log of the likelihood (L) of the data given the model. L is given by the joint binomial distribution:

$$L = \prod_{j=1}^{J} \binom{F_j}{f_j} p_j^{f_j}(1 - p_j)^{F_j - f_j} \tag{5}$$

where J is the number of inferences (that is, 4), f_j is the frequency with which an inference is endorsed, F_j is the total number of responses(that is, N), and p_j is the probability of drawing an inference according to our probabilistic model. To estimate the best-fitting parameter values, we minimised the log of (5), using a steepest descent search implemented in Mathematica's (Wolfram, 1991) *MultiStartMin* function (Loehle, 2000), which supplements the Newton–Raphson method with a grid-search procedure to ensure a global minimum. The log-likelihood ratio test statistic G^2, which is asymptotically distributed as χ^2, was used to assess the goodness of fit (Read & Cressie, 1988). This statistic evaluates the model fit by comparing the predicted values to a saturated model where all the values of p_j are set to the empirically observed proportions of cards selected. Within each condition in Byrne's Experiment 1, $P(p)$, $P(q)$ and ε were estimated from the data. As there were four data points and three parameters, G^2 was assessed against one degree of freedom. For model comparisons, the conventional 5 per cent level of significance is regarded as unreasonably large (Read & Cressie, 1988). The level of significance for rejection was therefore set at the 1 per cent level. The model could not be rejected for any of the three conditions (the best-fit parameter values are shown in parentheses), simple: $G^2(1) = .04$,

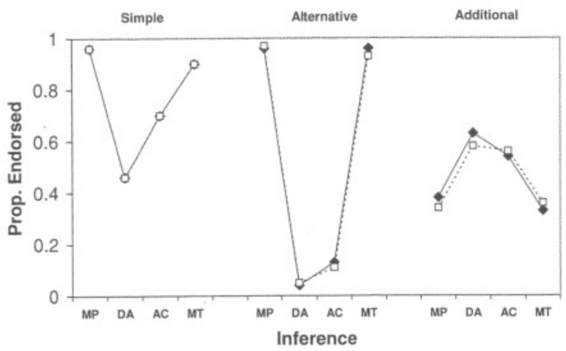

Figure 6.2 Fit between the model and Byrne's (1989: Experiment 1) suppression data, showing the probability of endorsing each inference observed (data) and predicted by Oaksford et al.'s (2000) probabilistic model (model)

$p > .50$ ($P(p) = .57$, $P(q) = .78$, $\varepsilon = .04$); alternative: $G^2(1) = .40$, $p > .50$ ($P(p) = .11$, $P(q) = .95$, $\varepsilon = .03$); or additional: $G^2(1) = .38$, $p > .50$ ($P(p) = .61$, $P(q) = .37$, $\varepsilon = .66$). Moreover, collapsing across conditions, the model could also not be rejected: $G^2(3) = .82$, $p > .50$.

According to probability theory, the conditional probabilities, $P(\text{not-}q \mid p)$ and $P(q \mid \text{not-}p)$, are independent of $P(p)$ and $P(q)$. However, according to our probability model, manipulating alternative and additional antecedents leads to changes in people's assessments of $P(p)$ and $P(q)$ (see Equations 2–4), as can be seen in the best-fit parameter values. It is important therefore that the changes in these parameters in response to manipulations of alternative and additional antecedents make intuitive sense. Let us look first at alternative antecedents. There are two cases to consider. First, how should being told that there are many different ways of starting a car affect one's assessment of the probability of its starting? Intuitively, it seems it should raise the probability: the more ways there are of making something happen, the greater the probability that it does happen. This common-sense principle seems to be based on the fact that people interpret events as part of a causal nexus rather than as isolated and independent. Second, how should being told that there are many different ways of starting a car affect one's assessment of the probability that the key is turned? Intuitively, it seems it should lower the probability: the more ways there are of making something happen, the lower the probability that any particular way need be invoked (but see below). These intuitions seem to be captured by Oaksford et al.'s (2000) model. In the simple condition in Experiment 1, the best fitting value for $P(p)$ was .57 and for $P(q)$ it was .78, but in the alternative antecedents condition, $P(p)$ fell to .11 and $P(q)$ increased to .95.

We now look at the effects of additional antecedents. Again, there are two cases to consider. First, how should being told that there are many different factors affecting whether a car will start affect one's assessment of the probability that the key is turned? Intuitively, it seems it should leave this probability unaffected: knowing that other factors are required for an action to succeed does not affect the probability that that action is performed. Second, how should being told that there are many different factors affecting whether a car will start affect one's assessment of the probability of its starting? Intuitively, it seems it should lower the probability: the more things that can prevent an action from producing the desired effect, the less likely that effect is to occur. Again, these intuitions seemed to be captured by Oaksford et al.'s (2000) model. In the additional antecedents condition, $P(p)$ stayed roughly the same at .61, while $P(q)$ decreased to .37. The suggestion that the model's behaviour may be consistent with various common-sense principles of reasoning is a prediction that is open to experimental test.

Modelling Suppression Effects II: Cummins et al. (1991)

Denise Cummins very kindly provided us with the raw data from her 1991 paper on suppression effects. In that experiment, a variety of causal conditionals were pre-tested for number of alternative and additional antecedents. These factors were also fully crossed in a 2×2 design; that is, rules were used that had many alternative and many additional antecedents (MM), many alternative and few additional antecedents (MF), few alternative and many additional antecedents (FM), and few alternative and few additional antecedents (FF). Participants rated each inference for each rule on a 6-point scale (1–6). To model these data, we re-scaled these ratings into the 0–1 probability scale by taking away 1 and dividing by 5. To fit the model to the data using the same procedure as above, we then multiplied each proportion by the sample size (27). This number was then rounded to the nearest integer, to obtain the frequency of participants endorsing an inference. We then fitted the model to the data, for each rule-type (MM, MF, FM and FF), as we described above. The model could not be rejected for any of the four rules (the best-fit parameter values are shown in parentheses), MM: $G^2(1) = 2.14$, $p > .10$ ($P(p) = .48$, $P(q) = .66$, $\varepsilon = .29$); MF: $G^2(1) = 5.58$, $p > .01$ ($P(p) = .49$, $P(q) = .67$, $\varepsilon = 25$); FM: $G^2(1) = .92$, $p > .20$ ($P(p) = .52$, $P(q) = .64$, $\varepsilon = .23$), or FF: $G^2(1) = 1.66$, $p > .10$ ($P(p) = .46$, $P(q) = .63$, $\varepsilon = .16$). Collapsing across conditions, the model could also not be rejected: $G^2(4) = 10.30$, $p > .02$. Consequently, as for Byrne (1989), the model provides good fits to Cummins et al.'s (1991) results.

We also fitted the model to each individual participant's results. We did this by first re-scaling the ratings as above, and then minimising the sum of squared differences between the data and the model, as in Oaksford et al. (2000). This procedure meant that best-fitting parameter values could be calculated for each rule type for each participant. These parameters could then be analysed statistically. First, however, we assessed the goodness of fit. We used the best-fit parameter values to calculate the probability of drawing each inference for each participant. We then compared the mean of these values to the re-scaled data means. We illustrate the fit in Figure 6.3, which shows the mean probabilities of each inference calculated from the data and calculated from the model. The error bars show the 95 per cent confidence intervals for the data. For this comparison, $R^2 = .84$, a result which shows that the model captures the general trend in the data quite well. We also computed the root mean squared scaled deviation (RMSSD) (Schunn & Wallach, 2001), which provides a scale-invariant estimate of how much the model diverges from the exact location of the data points in standard error units. RMSSD = 1.62, which means that, on average, the model deviated from the data values by only 1.62 standard error units. Thus, the model provides a good fit to the location of each data point.

Having established that the model provided a good fit to the data, we then asked whether the model's parameters vary in the way the model predicts. The model may fit the data

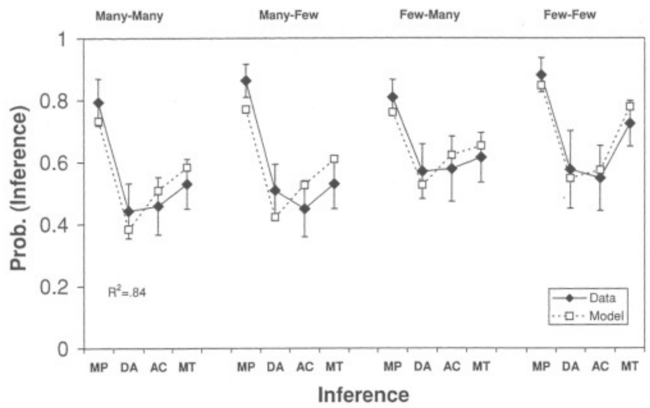

Figure 6.3 Fit between the model and Cummins et al.'s (1991) suppression data, showing the mean probability of endorsing each inference observed (Data) and mean value predicted by Oaksford et al.'s (2000) probabilistic model (Model). Error bars show 95% confidence intervals

very well, but not necessarily for the right reasons. For example, our model predicts that variation in additional antecedents should affect ε, the exceptions parameter. When there are many additional antecedents, ε should be higher than when there are few additional antecedents. Figure 6.1 also shows that there should be a relationship between ε and the number of alternative antecedents. As the number of exceptions increases, the probability that DA should be endorsed falls, and so the probability of alternative antecedents (DA′) rises. So, when there are many alternative antecedents, ε should be higher than when there are few alternative antecedents.

We analysed the best-fitting values of ε in a 2 × 2, within-subjects ANOVA with alternative (many vs. few) and additional (many vs. few) antecedents as factors (see Table 6.2). There was a significant main effect of additional antecedents: $F(1, 26) = 18.72$, $MSE = .006$, $p < .0005$. ε was significantly higher when there were many (mean = .25, SD = .16) than when there were few (mean = .19, SD = .13) additional antecedents. There was also a significant main effect of alternative antecedents: $F(1, 26) = 7.44$, $MSE = .010$, $p < .025$. ε was significantly higher when there were many (mean = .25, SD = .16) than when there were few (mean = .20, SD = .13) alternative antecedents. These results confirm the predictions of our model.

There was also a very close to significant interaction effect: $F(1, 26) = 4.19$, $MSE = .004$, $p = .051$. Simple effects comparisons showed that although alternative antecedents had a significant effect on ε when there were few additional antecedents—$F(1, 26) = 14.44$, $MSE = .006$, $p < .001$—they did not have a significant effect on ε when there were many additional antecedents: $F(1, 26) = 1.24$, $MSE = .008$, $p = .275$. According to our model, it would seem that when there are many additional antecedents, alternative antecedents must influence inference via $P(p)$ and $P(q)$ rather than ε. The other simple effects comparisons revealed that additional antecedents significantly influenced ε at both levels of alternative antecedents (few: $F(1, 26) = 17.02$, $MSE = .006$, $p < .001$, many: $F(1, 26) = 5.44$, $MSE = .004$, $p < .05$), although the effect was weaker when there were many alternative antecedents.

Table 6.2 Mean values of the best-fit parameters (ε, $P(p)$ and $P(q)$) for Cummins et al.'s (1991) experiment by alternative antecedents (Alts.: many vs. few) and additional antecedents (Adds.: many vs. few)

	Many Alts.				Few Alts.			
	Many Adds.		Few Adds.		Many Adds.		Few Adds.	
Parameter	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ε	.27	.18	.23	.13	.24	.13	.15	.12
$P(p)$	.47	.19	.45	.19	.49	.20	.41	.22
$P(q)$	.68	.15	.67	.15	.62	.16	.63	.21

We also analysed the $P(p)$ and $P(q)$ parameters (see Table 6.2). There are a couple of predictions about how these parameters should vary. In modelling Byrne's (1989) results above, we found quite marked changes in these parameters. In the best-fit parameter values, many alternative antecedents led to increases in $P(q)$ and decreases in $P(p)$, whereas many additional antecedents led to decreases in $P(q)$ but no changes in $P(p)$. We motivated these changes by appeal to various common-sense principles. We tested to what extent these patterns are present in Cummins et al.'s (1991) results. However, Byrne's (1989) explicit manipulation produced stronger effects (compare Figures 6.2 and 6.3), and she did not use a fully crossed design, so possible interaction effects could not be assessed. Consequently, any effects are likely to be weaker. Moreover, the interaction for ε suggests that the predicted differences for alternative antecedents will be seen only when there are many additional antecedents.

We analysed the best-fitting values of $P(p)$ and $P(q)$ in a 2 × 2 × 2, within-subjects ANOVA with parameter ($P(p)$ vs. $P(q)$), alternative antecedents (many vs. few) and additional antecedents (many vs. few) as factors (see Table 6.2). There was a significant main effect of parameter: $F(1, 26) = 14.71$, $MSE = 145$, $p < .001$. $P(q)$ (mean = .65, SD = .17) was significantly higher than $P(p)$ (mean = .45, SD = .20). This simply reflects the general constraint on our model that $P(q) > P(p)(1 - ε)$, so that when $ε = 0$, $P(q) > P(p)$. There was also a significant two-way interaction between parameter and alternative antecedents, $F(1, 26) = 4.88$, $MSE = .006$, $p < .05$, which was modified by a three-way interaction: $F(1, 26) = 4.67$, $MSE = .005$, $p < .05$. The three-way interaction partly reflects the prediction that the differences for alternative antecedents will be seen only when there are many additional antecedents. When this is the case, $P(q)$ was significantly higher—planned contrast: $F(1, 26) = 10.12$, $MSE = .006$, $p < .005$—when there were many alternative antecedents than when there were few alternative antecedents (see Table 6.2). Moreover, when there were many additional antecedents, there was a trend, although not significant, for $P(p)$ such that it was lower when there were many alternative antecedents than when there were few (see Table 6.2). These results are consistent with the changes in these parameter values observed in Byrne (1989), which we suggested may reflect various common-sense principles of reasoning.

However, there were other changes in parameters reflected in the three-way interaction. Specifically, when there were few alternatives, $P(p)$ was higher when there were many additional antecedents than when there were few. This difference was significant in a planned contrast: $F(1, 26) = 14.28$, $MSE = .005$, $p < .001$. This means, for example, that if you know a car can be started only by turning the key (few alternative antecedents), then the more conditions that need to be satisfied for turning the key to start the car (many additional antecedents), the more likely you are to turn the key (higher $P(p)$). The only intuitive motivation we can think of for such a principle is that, the more conditions that might need to be checked before taking an action, the more likely someone is not to bother and perform the action regardless, just to see if it works. Such a principle may be of limited generality.

In this section, we have shown how our model can provide detailed fits to the data on the suppression effect. Moreover, we have demonstrated that to achieve the fits the best-fit parameter values behave pretty much as would be expected. In the following sections, we explore some further suppression effects in conditional inference that can also be explained by this model.

Suppression Effects: Further Findings

In this section, we address a range of further findings on the suppression effect that would appear to be compatible with Oaksford et al.'s (2000) probabilistic model. We first look at the possible consequences of experiments such as those reported by Cummins et al. (1991), and Thompson (1994), where information about alternative and additional antecedents is provided implicitly.

Implicit Presentation

Cummins et al. (1991), Cummins (1995) and Thompson (1994) report results that were very similar to Byrne's (1989). However, they left information about additional and alternative antecedents implicit. That is, unlike Byrne (1989), these authors pre-tested rules for how many alternative and additional antecedents they allowed, and used these rules in the experimental task with no further explicit cueing as to the relevance of alternative and additional antecedents. These results seem directly to contradict Byrne, Espino and Santamaria (1999, p. 369), who have recently argued that people do *not* have, "a general insight into the idea that there may be alternatives or additional background conditions that are relevant to inferences". Byrne et al. (1999) appear to be arguing that, in the absence of explicitly provided information about alternative or additional antecedents, people do not retrieve it from long-term memory of world knowledge to decide whether to draw an inference. This view is not consistent with Cummins et al. (1991) or Thompson (1994), where, even when such information was left implicit, suppression effects were still observed. Consequently, participants must be accessing appropriate world knowledge to determine the likelihood that an inference can be drawn. This conclusion is further supported by the recent results of Liu et al. (1996). In their "reduced" inference condition, they presented participants with contentful material but without an explicit conditional premise; for example *knowing that the key has been turned, how probable is it that the car starts*? They found similar suppression effects as when they provided an explicit conditional premise. As Liu et al. (1996) argue, in the reduced inference condition, participants *must* be basing their inferences on accessing prior knowledge. *Pace* Byrne et al. (1999), that similar suppression effects were observed means that information about additional and alternative antecedents was being implicitly accessed.

Facilitating DA and AC

Byrne (1989) also showed that suppression effects can be removed by providing more information in the categorical premise. For example, given *if p then q* and *if r then q*, participants would be given the categorical premise, *p and r*. In Experiment 2, Byrne found that using materials like this removed all suppression effects. Indeed, using this manipulation produced a facilitation effect for DA and AC. We can explain this effect by the different ways additional and alternative antecedents affect the appropriate conditional probabilities. Only the number of alternative antecedents independently affects these probabilities, whereas the number of additional antecedents does not. For example, take the rule *if the key is turned, the car starts*. There are many exceptions to this rule: the car will not start if there is no petrol (*if there is petrol in the tank, the car starts*), if the battery is flat (*if the battery is charged, the car starts*) and so on. To make the MP inference in the first place, one must assume that all these possible additional conditions are *jointly* satisfied. Consequently, being told that *the key is turned and the battery is charged* is not going to affect people's estimate of the probability of MP, nor, by parity of reasoning, that of MT, as they have already assumed that this jointly necessary condition applies. Conversely, there are other ways to start cars, hot-wiring (*if hot-wired, the car starts*), jump-starting (*if jump-started, the car starts*) and so on. Each is individually sufficient to start the car; consequently, the more that are ruled out, the less likely the car is to start. Consequently, being told that *the key was not turned and the car was not hot-wired*, will increase someone's estimate of the probability of DA, and, by parity of reasoning, that of AC. In terms of Oaksford et al.'s model, this means that this manipulation decreases the probability that the car starts even though the key has not been turned. This explanation accounts for the facilitation effect for DA and AC that Byrne (1989) observed in her Experiment 2.

Graded Suppression of MP and MT

Stevenson and Over (1995) have shown variation in MP and MT inferences by concentrating not on the number of additional antecedents but on their likelihood. So participants could be told that,

If the key is turned, the car starts, (1)
If the battery is charged, the car starts, and that (2)
The battery is always (almost always, sometimes, rarely, very rarely) charged (3)

Participants are then given the categorical premise *the key is turned*. The manipulation in the third premise directly manipulates the likelihood of an exception; that is, ε ($P(\text{not-}q \mid p)$). Participants' willingness to endorse MP and MT tracked this manipulation, as would be predicted by our probabilistic account. If we look at Equations 1 and 4, for MP and MT, it is clear that as ε increases, the probability of MP decreases. A similar effect is also predicted for MT, although, if a and b are kept constant, the slope for MT will be steeper than for MP (see Figure 6.1). Interestingly, in Stevenson and Over's data, using the *always* instruction leads to a facilitation effect compared to the condition in which premise 3 is absent. This is because, although in premise 2 participants are told that further conditions need to apply to make the inference in premise 1, they are then told in premise 3 that these conditions always apply!

Chan and Chua (1994) used a similar manipulation to Stevenson and Over (1995) to reveal graded suppression of MP and MT. However, rather than use a further premise, like 3 above, they used different additional antecedents that varied in their relative salience for achieving the conclusion. For example, premise 2 is quite salient to whether the car starts or not. However, other less salient conditions can be imagined:

If the engine has not been removed overnight, the car will start, or (2′)
If it was not foggy last night, the car will start (2″)
(damp points can prevent ignition)

Chan and Chua (1994) observed the same graded suppression of MP and MT as observed by Stevenson and Over (1995).

George (1997) has also investigated graded suppression effects in conditional inference by directly introducing information in the conditional about the probability of the consequent given the antecedent. For example, he used rules such as *if Pierre is in the kitchen, it is (not) very probable that Marie is in the garden*. This manipulation directly affects ε; when the consequent includes "very probable", ε is low, and when it includes "not very probable", ε is high. Predictably, in his Experiment 1, there were more MP inferences in the very probable than in the not very probable condition.

Sufficiency and Suppressing DA and AC

George (1997) also found suppression effects for DA and AC when perceived sufficiency was reduced (for "valid arguments", see George, 1997, Table 6.1); that is, when ε is high. These effects are not predicted by a simple model based on the effects of additional and alternative antecedents. However, they are predicted by our probabilistic model. Examining Equations 2 and 3, for DA and AC, reveals that increasing ε while keeping a and b constant will also lead to reductions in the relevant conditional probabilities (see Figure 6.1). Intuitively this also makes sense because the numbers of exceptions and alternatives are related. For example, the rule *if the key is turned, the car starts* captures the *normal* and *most reliable* way (ε is as low as it can get) of starting cars. Alternative methods of starting cars are generally less reliable; for example, *if the car is bump-started, the car starts*, relies on further factors such as the speed being sufficiently high when you take your foot off the clutch and so on. So this alternative way of starting a car is also a less reliable way of starting a car; that is, ε is higher. Our probabilistic model captures this intuition and so can explain the suppression of DA and AC when perceived sufficiency is low.

Summary

Oaksford et al.'s simple probabilistic model appears to be capable of accounting for many of the suppression and facilitation effects in conditional inference. The key factor is that people's prior knowledge can be interpreted as affecting the subjective probabilities assigned to events or to their occurrence conditional on other events occurring. Most of our arguments for how the model explains these effects result from showing why, intuitively, a particular

manipulation should affect the relevant probabilities in a way that is consistent with our simple probabilistic model. In the next section, we see whether this pattern of explanation can be extended to order effects in conditional inference.

ORDER EFFECTS

Other important determinants of performance on conditional inference are the order of clauses within the conditional premise, that is, the difference between *if p then q* and *q only if p*, and the order of presentation of the premises and conclusion of a conditional inference. These manipulations are important and interesting, and we argue that they may also be amenable to a probabilistic treatment.

Clause Order

There are two consistent effects of the change of clause order. First, *the car starts only if the key is turned* leads to more AC and MT inferences and fewer MP and DA inferences than *if the key is turned, the car starts* (Evans, 1977; Roberge, 1978; Evans & Beck, 1981). Second, it has been found that paraphrasing a rule *if p then q*, using *only if* depends on two factors: temporal precedence, that is, which of p and q occurs first, and perceived necessity, that is, is p necessary for q or q necessary for p? (Evans, 1977; Evans & Newstead, 1977; Evans & Beck, 1981; Thompson & Mann, 1995). Note that in our example, turning the key (p) both precedes and is causally necessary for the car to start, and hence is best paraphrased as *q only if p*; the opposite paraphrase, *the key is turned only if the car starts*, seems pragmatically infelicitous. Thompson and Mann (1995) have also observed that these effects seem independent of content domain; that is, they occur for conditionals expressing causes, permissions, definitions or co-occurrence relations.

Our model clearly does not address the psycholinguistic findings on paraphrasing. However, we can look to see whether the principal effect of the *q only if p* rule revealed by these results, to emphasise the necessity of p for q, is amenable to a probabilistic treatment. Probabilistically, this effect seems to correspond to a situation where the probability of the consequent given that the antecedent has not occurred ($P[q \mid \text{not-}p]$) is lowered; that is, there are fewer alternative ways of achieving the effect. This immediately gives rise to the problem that if this probability decreases, the probability of the DA inference increases. But the observation in this literature is that DA and MP inferences decrease while AC and MT increase (Evans, 1977; Evans & Beck, 1981). However, an alternative interpretation is that the *q only if p* rule lowers the joint probability of q and not-p ($P[q, \text{not-}p]$). Under these circumstances, it is possible to rearrange the marginals, that is, $P(p)$ and $P(q)$, so that the probability of DA falls and the probability of AC rises. However, to model the increase in MT requires a fall in ε. This predicts an increase in MP inferences. But, in the data, the move to the *q only if p* rule leads to decreases in the frequency of MP endorsements, not increases. Consequently, however one tries to capture the effect of the *q only if p* rule, it seems that the model is bound to get the direction of change wrong for at least one inference. This is further borne out in fitting the model to Evans' (1977) results. The model could not be rejected for the *if p then q* rule, $G^2(1, N = 16) = 4.92, p > .02$, or for the *q only if p* rule, $G^2(1, N = 16) = 1.02, p > .30$ (see Figure 6.4). Moreover, although the predicted proportion

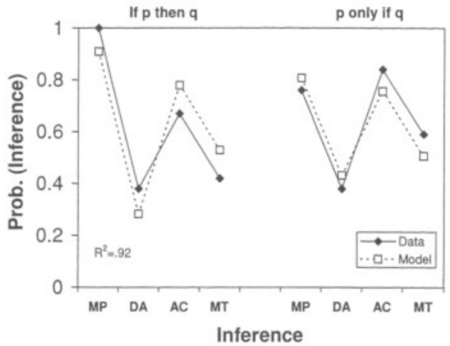

Figure 6.4 The observed and predicted probability of endorsements of each inference for the *if... then* and *only if* rules in Evans (1977)

of endorsements minimised the differences between rules, the direction of change was correct for MP, AC and MT. However, the model predicted an increase in DA inferences for the *q only if p* rule where either no change (Evans, 1977) or decreases were found (Evans & Beck, 1981).

Clause Order, Utilities and Conversational Pragmatics

It would appear that a straightforward probabilistic account of clause order effects is not available in our simple model. However, we believe that this may be because order manipulations have pragmatic functions that may be better captured via changes in utilities rather than probabilities. Oaksford, Chater and Grainger (1999) made a similar suggestion in the selection task. Order effects are usually discussed in terms of their direct effect on the construction of mental representations for the premises (Johnson-Laird & Byrne, 1991; Girotto et al., 1997; Evans et al., 1998). However, ordering of information may have pragmatic effects other than affecting the order in which a discourse representation is assembled. For example, order typically encodes the *topic* or focus of a discourse. For example, in an active sentence, the subject is the topic—hence the subject is mentioned first—whereas in a passive sentence the object is the topic—hence the object is mentioned first. Consequently, in interpreting even the relatively restricted discourse provided by a conditional syllogism, it is important to understand the normal communicative function of different sentential and clausal orderings.

Ordering manipulations may change the topic of a discourse. Changing the clausal ordering from *if p then q* to *q only if p* switches the topic from *p* to *q*. Communicative functions can be revealed by posing questions where one or the other linguistic form would be the most appropriate reply. For example, in response to the query, "What happens if I turn the key?", one might respond, "If you turn the key, the car starts." However, in response to the query, "Why did the car start?", one might reply, "Well, that happens only if you turn the key." Switching responses to these two queries would be pragmatically infelicitous. These examples suggest that the pragmatic function of *if p then q* is to focus attention on

what can be predicted given that *p* is known, whereas the pragmatic function of *q only if p* is to focus attention on explaining why *q* happened. Assuming that *p* temporally precedes *q*, as is normal (Comrie, 1986), MP and DA are the predictive inferences that are suppressed for the *q only if p* form, and AC and MT are the explanatory inferences that are facilitated for this form. This pattern of suppression and facilitation is consistent with the explanatory function of the *q only if p* form of the rule. The effect here is not to alter the relevant probabilities, but rather to alter the importance attached to the different inferences. Consequently, we argue that the utility to a reasoner of the different classes of inference is altered by the clause order manipulation. For the *q only if p* rule, people assign higher utility to the explanatory, AC and MT inferences, and a lower utility to the predictive, MP and DA inferences.

Premise and Conclusion Order

A further manipulation of order involves altering the order of premises and conclusion (Girotto, Mazzocco & Tasso, 1997; Evans, Handley & Buck, 1998). Girotto et al. (1997) showed that people are more willing to draw MT when the premises are presented in order (PCR) rather than in the standard order (PCS):

(PCR) The car has not started
If the key is turned, the car starts
(The key was not turned (C))

(PCS) If the key is turned, the car starts
The car has not started
(The key was not turned (C))

(The labels derive from Evans et al. (1998): "PC" means the conclusion (C) comes after the premises (P); "S" means that the premises are in standard order, conditional before categorical premise; and "R" means that this premise order is reversed.) The conclusion (in parentheses) was not presented in Girotto et al. (1997), as they used a production task where participants must spontaneously produce a conclusion rather than evaluate the validity of particular conclusions provided by the experimenter. Girotto et al. (1997) found no effect of this manipulation on MP, DA or AC, nor does it affect *q only if p* conditionals or biconditionals.

Evans, Handley and Buck (1998) used a similar manipulation but also varied the position of the conclusion. They used the orders in PCR and PCS (including the conclusion because Evans et al. used an evaluation task) and the following orders:

(CPR) The key was not turned (C)
The car has not started
If the key is turned, the car starts

(CPS) The key was not turned (C)
If the key is turned, the car starts
The car has not started

Evans et al. (1998) failed to replicate Girotto's et al.'s (1997) finding of increases in MT inferences with premise order reversals (PCR and PCS). What they did find was that both presenting the conclusion first (CPR and CPS vs. PCS) and reversing the premises (PCR and CPR vs. PCS) led to a reduction in negative conclusion bias, as we discussed in the introduction to this chapter. It was possible for Evans et al. (1998) to discover this because they used all the four conditions fully crossed with Evans' (1972) negations paradigm, where the four rules, *if p then q*, *if p then not-q*, *if not-p then q*, and *if not-p then not-q*, are presented. Evans et al.'s (1998) failure to replicate Girotto et al. (1997) suggests that it

Table 6.3 Mean values of the best-fit parameters ($P(p)$ and $P(q)$) for Evans et al.'s (1998) Experiment 2 by premise and conclusion order (PC vs. CP), premise order (standard vs. reverse), and whether the clause was negated or affirmative

	PC				CP			
	Standard		Reverse		Standard		Reverse	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Negated	.85	.07	.78	.05	.82	.04	.67	.04
Affirmative	.59	.10	.73	.01	.76	.01	.73	.11

may be too simplistic to attempt an interpretation that concentrates only on variation in MT inferences. We therefore concentrate on the effects described by Evans et al.

Probabilities and Conversational Pragmatics

Oaksford and Chater (1995a) suggested that conversational pragmatics may influence reasoning by affecting subjective probabilities. In that paper, we were concerned with accounting for the effects of relevance manipulations in Wason's selection task (Sperber, Cara & Girotto, 1995). We suggested that increasing the relevance of particular instances of a rule may result in increasing the degree of belief that such instances exist. Consequently, order effects that alter the topic of a sentence may also serve to increase the relevance and hence the subjective probability of a described property or event. This reasoning suggests investigating the probabilistic effects of ordering manipulations. We have done this by fitting our probabilistic model to the results from Evans et al.'s (1998) Experiment 2 that used the fully crossed design outlined in PCR to CPS. We fitted the model to the data for each of the four rules in the negations paradigm in each of the four conditions. We allowed all three parameters (a, b and ε) of the model to vary. The model could not be rejected for any of the 16 rule-condition pairs: average $G^2(1, N = 20) = 1.54$ (SD = 1.65), $p > .20$. Moreover, overall, the model could not be rejected: $G^2(16) = 24.63$, $p > .05$.

We investigated whether there were significant changes in parameter values dependent on the experimental manipulations. We first looked at the values of a and b, that is, of the probability of the antecedent ($P(p)$) and consequent ($P(q)$), respectively. To do this, we treated the parameters as our unit of analysis. $P(p)$ and $P(q)$ were estimated for each of the four rules in each of the four conditions, PCR to CPS.[1] Table 6.3 shows the mean values of the best-fit parameter values split by premise and conclusion order (PC vs. CP), premise order (S vs. R) and negation, that is, whether the parameter corresponds to a negated or an affirmative clause. According to Oaksford et al.'s (2000) model of the negations paradigm, negated clauses should correspond to high-probability categories. For example, our model predicts that $P(q)$ should be higher for the rule *if p then not-q* than for *if p then q*.

[1] In the analyses for the affirmative group the $P(q)$ value for the HL rule was omitted. This was because this rule is pragmatically odd. For example, it is like suggesting that *if something is black, it is a raven*, a statement known to be false because there are far more black things besides ravens. In Oaksford et al.'s (2000) model, the best fit values of $P(q)$ value for this rule were always high, and that is why we removed them from the analysis. However, this left too few data points in each cell (3). Consequently, we added a further data point corresponding to the mean of the remaining data points.

Consequently, we should see an effect of negation: negated clauses should correspond to higher probabilities.

We conducted a 2 × 4, mixed ANOVA with condition as a between-subjects factor and negation as the within-subjects factor.[2] There were main effects of negation, $F(1, 12) = 23.72$, $MSE = .002$, $p < .0005$, and of condition, $F(3, 12) = 3.27$, $MSE = .005$, $p = .059$, which were both modified by a significant interaction, $F(3, 12) = 19.51$, $MSE = .002$, $p < .0001$. Simple effects comparisons revealed that the best-fit parameters were significantly higher when they corresponded to a negated rather than an affirmative clause only for the PCS condition, $F(1, 12) = 70.44$, $MSE = .002$, $p < .0001$. This analysis really just redescribes Evans et al.'s (1998) finding that these order manipulations remove negative conclusion bias, in terms of our probabilistic model. However, why this happens in the model is interesting.

The difference between affirmative and negated clauses has narrowed for the non-standard orders mainly because of *increases* in the probabilities of the affirmative clause. The simple effect comparison for the affirmative clauses was significant: $F(3, 21) = 7.17$, $MSE = .003$, $p < .002$. To analyse this effect further, we carried out a one-way ANOVA on just the affirmative data. In post hoc Newman–Keuls tests, all the non-standard conditions had affirmative clauses that had significantly higher probabilities than the standard PCS condition at the 5 per cent level. No other differences approached significance. A very different pattern of results was found for the negated clauses. Again the simple effect comparison was significant: $F(3, 21) = 8.94$, $MSE = .003$, $p < .001$. However, now in post hoc Newman–Keuls tests, there were no significant differences between the PCS, PCR and CPS conditions; that is, the probabilities of the negated clauses remained high. However, all three of these conditions had negated clauses that had significantly higher probabilities than the CPR condition at the 5 per cent level.[3]

How could these order manipulations lead to these changes in subjective probability? We have suggested that making some piece of information relevant (Sperber, Cara & Girotto, 1995) may increase the subjective probability of an event (Oaksford & Chater, 1995a), and that making something the topic of a discourse may have the same effect. We should not be surprised at such changes in our subjective probabilities because in inference they are all relative to context. So, for example, the probability of encountering a llama on the streets of Cardiff is extremely low by our estimation. However, the probability of such an encounter at London Zoo is far higher. In all but the standard PCS condition, instead of the rule, one of the antecedent or consequent clauses is the first sentence of the limited discourse provided by these argument forms. That is, one of the antecedent or consequent clauses acts as the topic of the discourse. This explains why the probabilities of the affirmative clause rise. No corresponding increases in the probabilities of negated clauses occur because they are already high. This seems to make sense. For example, the probability that people are *not* going to encounter a llama on the streets of Cardiff is already very high. Telling them that they are not going to encounter one, that is, making *not* encountering llamas the topic of some discourse, is unlikely to increase their subjective estimate of this event.

Another aspect of the ordering manipulation concerns the coherence of the resulting discourse. PCS, PCR and CPS all seem to be coherent, whereas CPR does not. We illustrate

[2] We used a repeated-measures ANOVA because each parameter-rule combination occurs in each condition; for example, $P(p)$ for the *if p then q* rule was estimated for each of the four conditions.

[3] We cannot offer any explanation for the decline in the probability of categories corresponding to negated clauses in the CPR conditions.

this point by providing discourse examples of these different orders. We use the MP argument form throughout.

PCS′ When there is heavy rain in the Welsh Marches, there are often floods along the river Severn. In summer 1999, there was heavy rain in the Welsh Marches. Towns along the river Severn were flooded for days.

Here the causal relation between rain in the Welsh Marches and flooding on the Severn is the topic. After the second sentence, the invited inference is clearly what is stated in the final sentence, that is, the conclusion. So the function of this ordering seems to be to consider the consequences of the truth of the relationship described in the conditional.

PCR′ In summer 1999, there was heavy rain in the Welsh Marches. When there is heavy rain in the Welsh Marches, there are often floods along the river Severn. That year, towns along the river Severn were flooded for days.

Here the topic is the heavy rain in the Welsh Marches in the summer of 1999. Introducing the causal relation in the second sentence clearly invites the reader to consider the consequences of this fact. So the function of this ordering seems to be to consider the consequences of the fact described in the first sentence, that is, of the topic of the discourse.

CPS′ In summer 1999, towns along the river Severn were flooded for days. When there is heavy rain in the Welsh Marches, there are often floods along the river Severn. There was heavy rain in the Welsh Marches that year.

Here the topic is clearly the flooding along the river Severn in the summer of 1999. Introducing the causal relation in the second sentence clearly invites the reader to consider possible explanations of this fact. So the function of this ordering seems to be to consider possible explanations of the fact described in the first sentence, that is, of the topic of the discourse.

Note that PCR′ and CPS′ are interchangeable; for example, if we presented AC in CPS form, the resulting discourse would be identical to MP in PCR form. This predicts that MP and AC inferences in both PCR′ and CPS′ should be endorsed at similar levels, whereas normally MP inferences are endorsed much more strongly. In Evans et al.'s (1998) results, the difference between endorsements of MP and AC was 21 per cent in the PCS condition, whereas in the PCR and CPS conditions it was only 5 per cent and 6 per cent, respectively (DA and MT inferences were at similar levels overall in all conditions).

?CPR′ In summer 1999, towns along the river Severn were flooded for days. There was heavy rain in the Welsh Marches that year. When there is heavy rain in the Welsh Marches, there are often floods along the river Severn.

We have put a question mark before CPR′ because, although the discourse is not nonsensical, it is not wholly coherent. The first sentence introduces a fact. For the second sentence to be coherent, it must be regarded as relevant to this fact. The only way it seems this can come about is when the fact in the second sentence is explanatory of (or can be predicted from) the first. That is, the second sentence is a tentative attempt to suggest a causal relation between the facts described in these juxtaposed sentences (see Comrie, 1986, on the use of sentential

juxtaposition to suggest causation). The final conditional sentence then states that there is such a relation. This seems to violate the pragmatic maxim of quantity (Levinson, 1983): a statement should be as informative as is required for the current discourse. The problem here is that the first two sentences in CPR′ suggest only that there *may* be a causal relation between these two facts. However, the final sentence makes the more informational statement that there actually is such a relationship. According to the maxim of quantity, if someone knew that such a relationship existed, there was no point in just suggesting that it did in the first two sentences. Rather, as in the other orders, this should be stated upfront.

What effects might we predict from this apparent violation of the maxim of quantity? The most obvious effect of stating only that a causal relation *may* exist in the first two sentences is to weaken participants' belief in the conditional describing that relation in the final sentence. Although we have argued that there is not always a relation between the degree to which a rule is believed and exceptions (Chater & Oaksford, 1999a), this is often the case. For example, your degree of belief that *if something is a widget, it is blue* would probably be severely reduced if you are told that *most widgets are not blue*. This suggests that the effect of CPR′ may be to increase the probability of exceptions; that is, this pragmatic account of order effects suggests that ε rises in the CPR′ condition. We therefore statistically compared the best-fit values of ε between conditions in a one-way, within-subjects ANOVA. The result was significant: $F(3, 9) = 5.57$, $MSE = .003$, $p < .025$. In post hoc Newman–Keuls tests, the CPR condition (mean = .21, SD = .08) had a higher mean probability of exceptions than all other conditions at the 5 per cent level (PCS: mean = .08, SD = .02; PCR: mean = .08, SD = .03; CPS: mean = .10, SD = .03). No other differences approached significance. Consequently, it would seem that Evans et al.'s (1998) results are consistent with the likely probabilistic effects of our pragmatic account of the premise-conclusion ordering manipulation.

Summary

We have argued that the pragmatic effects of ordering manipulations are compatible with Oaksford et al.'s (2000) probabilistic account of conditional inference. In the case of premise and conclusion order effects (Girotto et al., 1997; Evans et al., 1998), the explanation is mediated by the pragmatic effects of these manipulations. This is consistent with Oaksford and Chater's (1995a) arguments about the probabilistic effects of pragmatic phenomena. In explaining the effect of clause order changes (e.g., Evans, 1977; Roberge, 1978; Evans & Beck, 1981), we argued that a decision theoretic perspective may be required to capture the different explanatory and predictive functions of the *if p then q* and *q only if p* rule forms.

CONCLUSIONS

We have argued that a probabilistic approach can resolve many of the problems of logic-based approaches to non-monotonic or defeasible reasoning. These problems are revealed by phenomena such as the failure of strengthening of the antecedent for everyday conditionals. Adopting a probabilistic approach leads naturally to the expectation of suppression effects in conditional inference, which we modelled, using a simple contingency table approach to the meaning of conditional statements. We also showed how the same model accounts for a variety of other suppression and facilitation effects. We also looked at order effects. We argued that these phenomena can be explained in a rational, probabilistic framework as long as close attention is paid to the pragmatic function of these different ordering manipulations and their likely probabilistic effects. In sum, together with Oaksford et al.'s (2000) account of negative conclusion bias, we have shown that a rational probabilistic model may explain many of the major effects in the psychology of conditional inference. This is important because previous commentators in this area have seen these data as providing evidence of systematic bias or of the operation of suboptimal algorithms for conditional reasoning.

REFERENCES

Adams, E. (1966). Probability and the logic of conditionals. In J. Hintikka & P. Suppes (eds), *Aspects of Inductive Logic*. Amsterdam: North-Holland.

Adams, E. (1975). *The Logic of Conditionals: An Application of Probability to Deductive Logic*. Dordrecht: Reidel.

Anderson, A. R. & Belnap, N. D. (1975). *Entailment: The Logic of Relevance and Necessity*, vol. 1. Princeton, NJ: Princeton University Press.

Anderson, J. R. (1995). *Cognitive Psychology and Its Implications*. New York: W. H. Freeman.

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

Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 1–21.

Byrne, R. M. J., Espino, O. & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language*, *40*, 347–373.

Chan, D. & Chua, F. (1994). Suppression of valid inferences: Syntactic views, mental models, and relative salience. *Cognition*, *53*, 217–238.

Chater, N. & Oaksford, M. (1990). Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn. *Cognition*, *34*, 93–107.

Chater, N. & Oaksford, M. (1999a). Information gain and decision-theoretic approaches to data selection. *Psychological Review*, *106*, 223–227.

Chater, N. & Oaksford, M. (1999b). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, *38*, 191–258.

Chater, N. & Oaksford, M. (2000). The rational analysis of mind and behaviour. *Synthese*, *122*, 93–131.

Chater, N. & Oaksford, M. (2001). Human rationality and the psychology of reasoning: Where do we go from here? *British Journal of Psychology*, *92*, 193–216.

Clark, K. L. (1978). Negation as failure. In *Logic and Databases* (pp. 293–322). New York: Plenum Press.

Comrie, B. (1986). Conditionals: A typology. In E. C. Traugott, A. ter Meulen, J. S. Reilly & C. A. Ferguson (eds), *On Conditionals* (pp. 77–99). Cambridge: Cambridge University Press.

Cook, S. A. (1971). The complexity of theorem proving procedures. In *Proceedings of the Third Annual ACM Symposium on the Theory of Computing* (pp. 151–158). New York: Association for Computing Machinery.

Cummins, D. D. (1995). Naïve theories and causal deduction. *Memory and Cognition*, *23*, 646–658.

Cummins, D. D., Lubart, T., Alksnis, O. & Rist, R. (1991). Conditional reasoning and causation. *Memory and Cognition*, *19*, 274–282.

Evans, J. St. B. T. (1972). Interpretation and "matching bias" in a reasoning task. *Quarterly Journal of Experimental Psychology*, *24*, 193–199.

Evans, J. St. B. T. (1977). Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology*, *29*, 297–306.

Evans, J. St. B. T. & Beck, M. A. (1981). Directionality and temporal factors in conditional reasoning. *Current Psychological Research*, *1*, 111–120.

Evans, J. St. B. T. & Newstead, S. E. (1977). Language and reasoning: A study of temporal factors. *Cognition*, *8*, 265–283.

Evans, J. St. B. T., Handley, S. & Buck, E. (1998). Order effects in conditional reasoning. *British Journal of Psychology*, *89*, 383–403.

Evans, J. St. B. T., Newstead, S. E. & Byrne, R. M. J. (1993). *Human Reasoning*. Hillsdale, NJ: Erlbaum.

George, C. (1997). Reasoning from uncertain premises. *Thinking and Reasoning*, *3*, 161–190.

Ginsberg, M. L. (ed.) (1987). *Readings in Nonmonotonic Reasoning*. Los Altos, CA: Morgan Kaufmann.

Girotto, V., Mazzocco, A. & Tasso, A. (1997). The effect of premise order in conditional reasoning: A test of the mental model theory. *Cognition*, *63*, 1–28.

Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Kripke, S. (1963). Semantical considerations on modal logic. *Acta Philosophica Fennica*, *16*, 83–94.

Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, C. I. (1918). *A Survey of Symbolic Logic*. Berkeley, CA: California University Press.

Lewis, D. (1973). *Counterfactuals*. Oxford: Oxford University Press.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, *85*, 297–315.

Liu, I., Lo, K. & Wu, J. (1996). A probabilistic interpretation of "If-then". *Quarterly Journal of Experimental Psychology*, *49A*, 828–844.

Loehle, C. (2000). Global Optimization 4.0 [computer program]. Naperville, IL: Loehle Enterprises.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

McCarthy, J. M. (1980). Circumscription: A form of non-monotonic reasoning. *Artificial Intelligence*, *13*, 27–39.

McCarthy, J. M. & Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer & D. Michie (eds), *Machine Intelligence* (vol. 4.) New York: Elsevier.

McDermott, D. (1982). Non-monotonic logic II: Non-monotonic modal theories. *Journal of the Association for Computing Machinery*, *29*, 33–57.

McDermott, D. (1987). A critique of pure reason. *Computational Intelligence*, *3*, 151–160.

McDermott, D. & Doyle, J. (1980). Non-monotonic logic I. *Artificial Intelligence*, *13*, 41–72.

Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608–631.

Oaksford, M. & Chater, N. (1995a). Information gain explains relevance which explains the selection task. *Cognition*, *57*, 97–108.

Oaksford, M. & Chater, N. (1995b). Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning*, *1*, 121–152.

Oaksford, M. & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, *103*, 381–391.

Oaksford, M. & Chater, N. (1998a). *Rationality in an Uncertain World: Essays on the Cognitive Science of Human Reasoning*. Hove: Psychology Press.

Oaksford, M. & Chater, N. (1998b). A revised rational analysis of the selection task: Exceptions and sequential sampling (pp. 372–398). In M. Oaksford & N. Chater (eds), *Rational Models of Cognition*. Oxford: Oxford University Press.

Oaksford, M. & Stenning, K. (1992). Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 835–854.

Oaksford, M., Chater, N. & Grainger, B. (1999). Probabilistic effects in data selection. *Thinking and Reasoning*, *5*, 193–243.

Oaksford, M., Chater, N. & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 883–899.

Oaksford, M., Chater, N., Grainger, B. & Larkin, J. (1997). Optimal data selection in the reduced array selection task (RAST). *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 441–458.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufman.

Pylyshyn, Z. W. (ed.). (1987). *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex.

Read, T. R. C. & Cressie, N. A. C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Berlin: Springer-Verlag.

Reiter, R. (1985). On reasoning by default. In R. Brachman & H. Levesque (eds), *Readings in Knowledge Representation* (pp. 401–410). Los Altos, CA: Morgan Kaufmann.

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

Roberge, J. J. (1978). Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology*, *30*, 705–716.
Schunn, C. D., & Wallach, D. (2001). Evaluating goodness-of-fit in comparison of models to data. Unpublished manuscript. Learning Research and Development Center, University of Pittsburgh, US.
Shoam, Y. (1987). *Reasoning About Change*. Boston, MA: MIT Press.
Shoam, Y. (1988). Efficient reasoning about rich temporal domains. *Journal of Philosophical Logic*, *17*, 443–474.
Sperber, D., Cara, F. & Girotto, V. (1995). Relevance explains the selection task. *Cognition*, *57*, 31–95.
Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (ed.), *Studies in Logical Theory*. Oxford: Oxford University Press.
Stevenson, R. J. & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology*, *48A*, 613–643.
Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory and Cognition*, *22*, 742–758.
Thompson, V. A. & Mann, J. M. (1995). Perceived necessity explains the dissociation between logic and meaning: The case of "only if". *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 1554–1567.
Veltman, F. (1985). Logics for Conditionals. PhD Thesis, Faculteit der Wiskunde en Natuurwetenschappen, University of Amsterdam.
Wolfram, S. (1999). Mathematica 4.0 [computer program]. Cambridge: Cambridge University Press.

PART II

Judgment