*British Journal of Psychology* (1984), **75**, 451–468 *Printed in Great Britain*



# Heuristic and analytic processes in reasoning*

**Jonathan St B. T. Evans**

A general two-stage theory of human inference is proposed. A distinction is drawn between *heuristic* processes which select items of task information as 'relevant', and *analytic* processes which operate on the selected items to generate inferences or judgements.

These two stages are illustrated in a selective review of work on both deductive and statistical reasoning. Factors identified as contributing to heuristic selection include perceptual salience, linguistic suppositions and semantic associations. Analytic processes are considered to be context dependent: people reason from experience, not from inference rules. The paper includes discussion of the theory in comparison with other contemporary theories of human inference, and in relation to the current debate about human rationality.

Hundreds of experiments on human reasoning – deductive, inductive and statistical – have been reported in the literature (for recent books see Johnson-Laird & Wason, 1977; Nisbett & Ross, 1980; Evans, 1982*a*; Kahneman *et al.*, 1982). Two general features of the results are (1) that *observed* competence is often low, i.e. many inferential errors occur as measured by standard normative systems, and (2) that performance is highly inconsistent across situations, and susceptible to a host of experimental variables. These results, in turn, have led to a major debate about the implications for human rationality (see, for example, Cohen, 1981, and associated commentaries).

In this paper a general theoretical framework is proposed, within which to view the results of reasoning research. The motivation for this is firstly to show why reasoning errors are both common and inconsistent across situations, and secondly to provide a framework for the design of future research on reasoning. While it is not concerned with the proposal of models of deduction for specific tasks, the framework has clear implications for the manner in which such models should be constructed.

The theory involves a revision of some of my previous writing on the subject (e.g. Evans, 1982*a*, chapter 12) but a natural development of my discussion of selective processes in reasoning (Evans, 1983*a*). In the latter discussion it was suggested that errors of reasoning often occur because subjects fail to attend to the logically relevant information, or because they process information in a directional manner. One source of attentional bias discussed was simply the perceptual salience of the presented information. This paper is concerned with the explanation of more subtle effects – in particular one of the problems raised in discussion (Evans, 1982*b*) of Kahneman & Tversky (1982*a*). The problem is this: why is it that sometimes logically relevant information is highly salient in the presentation and yet apparently ignored by the subject? Does this necessarily imply that the subject is 'irrational' or in possession of faulty inference rules?

In answering these questions a distinction is drawn between two types of thought process which are termed *heuristic* and *analytic*. The function of the heuristic process is selection. The outcome of heuristic processing is a judgement of *relevance* about features of the problem. Information deemed 'irrelevant' is not processed further. 'Relevant' information is then subjected to analytic processing. The nature of these analytic processes is discussed in the final section of this paper. Suffice it to say at present that the function of analytic processes is to generate some form of inference or judgement from the information selected.

* This is an expanded version of a paper read to a conference on the 'Psychological content of logic', at the University of Tilburg, Holland, in October 1982.

I have long argued (e.g. Evans, 1972*a*) that reasoning experiments tend to be interpreted in a simplistic manner. Whilst those favouring a competence/performance distinction now recognize that an error of performance does not necessarily imply defective understanding of logic, they still tend to assume that correct performance *does* imply such understanding (e.g. Cohen, 1981). The implications of the present distinction for the rationality debate are discussed in a later section, but the essence is this: arguments for rationality should centre on the nature of analytic processes, but without an understanding of heuristic processes they are doomed to failure. The reason is quite simple: we cannot decide whether reasoning is rational unless we know what the subjects are reasoning *about*, and that is determined heuristically.

It is necessary to point out that the present use of the term 'heuristic' is quite different from that of Tversky & Kahneman (e.g. 1974). They use the term to refer to (conscious?) strategies which provide short-cut methods of solving inferential and judgemental problems. The term 'heuristic' in this paper refers to *pre-attentive* processes whose function is to select 'relevant' information for analytic processing. It determines what judgements are made about, rather than the way in which they are made.

An illustration of the types of thought that are involved is given by the example of selecting a move in chess. The aspect of human chess playing which has been so difficult to mimic in artificial intelligence programs is the heuristic selection of moves to be analysed (Newell & Simon, 1972). Strong players immediately 'see' that only a few moves are 'relevant' in the position. At each stage of the analysis – considering the opponent's reply, your reply to his, etc. – this 'automatic' reduction of the search space takes place (although *forced* replies can be computed in an analytic manner). These are the sort of heuristic processes that we are concerned with here – pre-attentive, rapid and indescribable by the person using them. The *choice* of alternatives in chess is, however, based on analytic processes, which evaluate the positions attained by computing material gain or, more mysteriously, judging positional advantage. Master chess players also use their analytic processes to annotate the games when they are published, in a manner analogous to *post hoc* justifications of logical decisions (see Wason & Evans, 1975, and below).

The argument advanced in this paper is, in fact, a natural development of the view that inferential behaviour should be regarded as a form of problem solving (see Newell, 1980; Evans, 1982*a*, chapter 11; Sternberg, 1982). Whilst I agree with Rips (1983*a*) that the theory of problem spaces is insufficient to account for reasoning, it does provide a useful framework for relating the study of reasoning to other types of thought. Thus analytic processes are seen to test information which has been generated selectively by heuristic processes. The present theory also supports Einhorn & Hogarth's (1981) discussion of 'figure-ground' effects in decision making.

The paper falls into two main sections. The first demonstrates the advantages of the heuristic/analytic distinction in explaining several examples of empirical reasoning research. The second takes up more general theoretical implications, including those for the rationality debate.

## Empirical studies

The object of this section is to provide empirical illustrations of the advantages of the heuristic/analytic distinction. It is clearly possible to consider only a very small portion of the literature in the space available, but some breadth is provided by the use of examples from the disjoint literatures on deductive and statistical inference.

### *Conditional reasoning*

In this section consideration is limited to two tasks designed to test people's understanding of the logic of conditionals: the *truth table task* and the *selection task*. The former task is a

rather specialized version of the general paradigm of sentence verification, in which most research has focused on simple assertions and denials of information that can be checked in a picture presented against knowledge stored in semantic memory. A comparatively small number of studies have extended this paradigm to the investigation of sentences like conditionals and disjunctives which connect two distinct propositions. Because of their obvious connections with propositional logic, such studies are regarded as investigating the *psychological truth tables* for linguistic 'rules'. For the present, our discussion will focus on the investigation of conditional sentences of the form 'if $p$ then $q$'. It is commonly assumed that such a task provides a 'direct' measure of the subjects' logical representation of the connective 'if–then' against which their reasoning accuracy with the sentence can be measured (e.g. Legrenzi, 1970; Marcus & Rips, 1979). As we shall see, this is an important conceptual error, and arises from the failure to recognize that the truth table task involves heuristic as well as analytic processes.

Wason (1966), among others, suggested that people have not two but *three* truth values for conditional sentences: true, false and *irrelevant*. In logical notation there are four cases to be considered: TT – both antecedent and consequent of the rule are true; TF – antecedent true, consequent false; FT – antecedent false, consequent true; FF – both antecedent and consequent false. In standard propositional logic, if the conditional were treated as material implication, then it would be false in the case TF, but otherwise true. Wason suggested that conditionals in natural language have a 'defective'* truth table, which differs by assigning the value irrelevant to the FT and FF cases. For example, the rule 'If it is a dog then it has four legs' is true for four-legged dogs, false for three-legged dogs, but *irrelevant* to tables and chairs, no matter how many legs they may have.

Johnson-Laird & Tagart (1969) found that most subjects did classify instances of conditional sentences in this way, from which one might infer that people do indeed 'possess' defective truth tables for conditional rules. Experiments introducing negative components into the sentences, however, showed this view to be too simple (Evans, 1972*b*, 1975; Evans & Newstead, 1977). In these experiments a factor other than the logical cases had a major influence on the distribution of 'irrelevant' judgements: the number of matches between items *named* in the rules and in the instances. The effect of introducing negatives into the rules is to rotate the presence of such matches across the four logical cases (see Table 1). For example, on the rule 'If the letter is B then the number is not 1', the TF case is produced by an instance B1 which is a double match. If the position of the negative is shifted as in 'If the letter is not B then the number is 1' then the TF case is produced by a double mismatch, e.g. J8.

Evans (1972*b*) using different (but equally artificial) materials asked subjects to *construct* all cases they could think of which verified and falsified the four forms of conditional rule shown in Table 1. Any logical case omitted was *inferred* by the experimenter to be irrelevant. The result was a 'matching bias': all else being equal, the subjects were more likely to select a case as either true or false if the values named in the instance matched those in the rule than if they did not (cases of one match and one mismatch were intermediate).

In itself, this could reflect a simple 'availability' bias or response priming effect (cf. Pollard, 1982). That is, the mention of the values in the rule increases their salience and biases the subject to choose them. What makes the phenomenon much more interesting is the discovery that an *evaluation* task procedure leads to an identical result (Evans, 1975). In this procedure all possible logical cases are presented to the subject in turn, thus ensuring the 'availability' of each. Subjects are given the choice of three classifications:

* The term 'defective' seems to imply that material implication is 'correct', which is hard to defend linguistically. I intend no such connotation but use the term only as a conventional label for a particular truth table.

**Table 1.** Examples of the four logical cases for each of the four conditional rules produced by permuting the presence of negative components

| Rule | Logical case | | | |
|---|---|---|---|---|
| | TT | TF | FT | FF |
| If $p$ then $q$ | $pq$ | $p\bar{q}$ | $\bar{p}q$ | $\bar{p}\bar{q}$ |
| *If the letter is B then the number is 4* | B4 | B5 | J4 | P7 |
| If $p$ then not-$q$ | $p\bar{q}$ | $pq$ | $\bar{p}\bar{q}$ | $\bar{p}q$ |
| *If the letter is J then the number is not 3* | J6 | J3 | P5 | W3 |
| If not-$p$ then $q$ | $\bar{p}q$ | $\bar{p}\bar{q}$ | $pq$ | $p\bar{q}$ |
| *If the letter is not T then the number is 8* | Z8 | B2 | T8 | T6 |
| If not-$p$ then not-$q$ | $\bar{p}\bar{q}$ | $\bar{p}q$ | $p\bar{q}$ | $pq$ |
| *If the letter is not D then the number is not 7* | L2 | R7 | D9 | D7 |

they can say that each instance conforms to the rule, contradicts the rule or is irrelevant to the rule. Subjects classify as 'irrelevant' precisely those cases which they would not have selected under the construction task procedure. Data from one such truth table evaluation experiment (Evans & Newstead, 1977) are shown in Table 2.

The parallel between the construction and evaluation tasks is very important and illustrates the difference between 'relevance' and 'availability'. The reason that subjects do not choose mismatching cases is not because they do not consider them, but because they do not consider them *relevant*. We now consider a specific explanation of the data shown in Table 2. The proposed underlying processes are illustrated in Fig. 1. First, there is a heuristic stage in which some instances are classed as irrelevant. *Only* the relevant ones go on to the second, analytic stage in which they are classified as true or false.

Evans & Newstead (1977) actually found that *two* factors significantly influenced the distribution of irrelevant responses. Careful examination of Table 2 reveals that such cases are more often associated with ($\bar{p}\bar{q}$) mismatching cases (top right to bottom left diagonal) than with ($pq$) matches (top left to bottom right). However, they are also far more often associated with FT and FF cases than with TT and TF cases. It is proposed that *both* factors reflect the operation of linguistically cued relevance judgements at the heuristic stage of the decision. False antecedent cases are liable to be rejected as irrelevant because the linguistic function of the conditional sentence is to direct attention to the supposition of the truth of the antecedent (Rips & Marcus, 1977). Matching bias is also linguistically cued, in that the use of a negative does not alter the fact that the proposition which it modifies is the *topic* of the sentence.

Consider the following:

(1) If I go by train I will not be tired when I arrive.
(2) If I do not go by train I will be tired when I arrive.

The topics of both sentences are the acts of travelling by train and of arriving tired. If asked to describe a situation which would show either rule to be false, people would tend in either case to suggest the combination of the two events, although the logical status of this is TF for (1) and FT for (2). The events specified – regardless of negatives – provide the linguistic topic, and are thus more likely to be coded as relevant by the subject.

This interpretation of the matching bias effect is supported by a recent experiment

**Table 2.** Percentage frequencies of truth table classifications for 'if–then' rules in the study of Evans & Newstead (1977)

| Rule and classification | Logical case TT | TF | FT | FF |
|---|---|---|---|---|
| If $p$ then $q$ | | | | |
| True | 100 | 3 | 0 | 3 |
| False | 0 | 97 | 63 | 9 |
| Irrelevant | 0 | 0 | 38 | 88 |
| | ($pq$) | ($p\bar{q}$) | ($\bar{p}q$) | ($\bar{p}\bar{q}$) |
| If $p$ then not-$q$ | | | | |
| True | 100 | 6 | 3 | 38 |
| False | 0 | 94 | 13 | 22 |
| Irrelevant | 0 | 0 | 84 | 41 |
| | ($p\bar{q}$) | ($pq$) | ($\bar{p}\bar{q}$) | ($\bar{p}q$) |
| If not-$p$ then $q$ | | | | |
| True | 94 | 6 | 3 | 31 |
| False | 3 | 59 | 75 | 28 |
| Irrelevant | 3 | 34 | 22 | 41 |
| | ($\bar{p}q$) | ($\bar{p}\bar{q}$) | ($pq$) | ($p\bar{q}$) |
| If not-$p$ then not-$q$ | | | | |
| True | 94 | 13 | 19 | 41 |
| False | 6 | 84 | 41 | 31 |
| Irrelevant | 0 | 3 | 41 | 28 |
| | ($\bar{p}\bar{q}$) | ($\bar{p}q$) | ($p\bar{q}$) | ($pq$) |

*Note.* Deviations of some column totals from 100 are due to rounding errors.

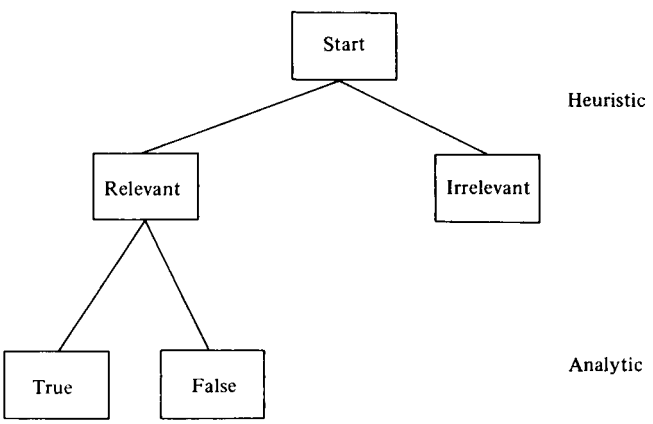



**Figure 1.** A two-stage model of the processes underlying performances on the truth table task.

reported by Evans (1983*b*). The instances to be evaluated are presented as conjunctive sentences. An example of a TF case is as follows:

Rule: If the letter is D then the number is 4.
Instance: The letter is D and the number is 7.

This produces one mismatch – in the consequent. The implicit negation of the 4 in the rule by the 7 in the instance is similar to that of all previous experiments. A second group was given instances where such negations occurred explicitly, for example:

Instance: The letter is D and the number is not 4.

In this way the letter and number named – and hence the linguistic topic of the instance – was always the same as that of the rule, regardless of the logical case studied. As predicted, this group showed substantial and significant reduction of matching bias as compared with the control group. In addition, they showed a significant increase in logically correct classifications of TT and TF cases (as 'true' and 'false' respectively). This is also consistent with the model, since the reduced elimination of such cases at the heuristic stage renders them more liable to analytic processing.

A different problem which also tests subjects' understanding of the truth conditions of conditional sentences is the selection task, first presented by Wason (1966). Many experiments have been performed with this problem, and a number of reviews and discussions of this research have appeared recently (Evans, 1982*a*, chapter 9; Pollard, 1982; Manktelow, 1982; Griggs, 1983; Wason, 1983). The coverage here is therefore restricted to that necessary to describe the application of the heuristic/analytic distinction to the explanation of selection task performance.

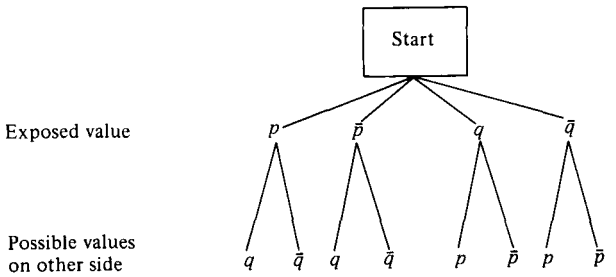
The selection task has been investigated with both artificial problem content – similar to that of the truth table tasks discussed above – and with thematic content. The artificial form is considered first. In a typical example the subjects are shown a set of cards, each of which has a letter on one side and a number on the other. They are then shown four such cards lying on a table with the exposed sides showing 'B', 'G', '2' and '9'. They then are given a rule which applies to the four cards and may be true or false:

If there is a B on one side of the card, then there is a 2 on the other side of the card.

The subject is then asked to select those cards, and *only* those cards, which it would be necessary to turn over in order to find out whether the rule is true or false. The most common answers are B or B and 2. In general terms, for a rule 'if $p$ then $q$', the choices are $p$ or $p,q$. Both are incorrect: the subject should choose B and 9 ($p$ and $\bar{q}$). In order to see why, examine the normative 'problem space' shown in Fig. 2.

To solve the problem subjects must construct the entire problem space shown in Fig. 2. That is, they must consider each card in turn, and each of two logical possibilities on the back. Then they must apply a truth table to evaluate all eight combinations. Finally, they must apply a decision rule to the effect that all and only cards which *might* produce a falsifying combination must be selected. This leads to the choice of $p$ and $\bar{q}$, assuming either material *or* defective implication (if the rule is treated as an equivalence all four cards should be selected).

Evidently most subjects deviate from this ideal in some way. Wason's (1966) original suggestion was that subjects use a wrong decision rule, i.e. they choose cards which could *verify*, rather than falsify the rule (in accordance with the defective truth table). Johnson-Laird & Wason (1970) proposed a model with several decision rules, and also a selective mechanism by which some subjects focus only on the cards $p$ and $q$. Neither of these theories can account for card selection when negatives are introduced into the rule, as in the studies by Evans & Lynch (1973) and Manktelow & Evans (1979, see Table 3). Card selections reflect exactly the same two factors which determine the *relevance* judgement on the truth table task (Table 2). That is to say there is a matching bias – $p$ and $q$ are chosen more often than $\bar{p}$ and $\bar{q}$ – and a bias over all rules to choose the true antecedent case (TA) most often, and the false antecedent (FA) least often.




*Evaluation*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Implication (full) | T | F | T | T | T | T | F | T |
| Implication (defective) | T | F | I | I | T | I | F | I |
| Equivalence (full) | T | F | F | T | T | F | F | T |

**Figure 2.** A 'problem space' for the Wason selection task, with evaluations based on three alternative truth tables. T = true; F = false; I = irrelevant.

**Table 3.** Percentage frequencies of selection of cards in Manktelow & Evans' (1979) Expts I and II combined (abstract rules, $n = 48$)

| | TA[a] | FA | TC | FC[a] |
|---|---|---|---|---|
| If $p$ then $q$ | ($p$) | ($\bar{p}$) | ($q$) | ($\bar{q}$) |
| | 89 | 15 | 65 | 29 |
| If $p$ then not-$q$ | ($p$) | ($\bar{p}$) | ($\bar{q}$) | ($q$) |
| | 98 | 6 | 27 | 75 |
| If not-$p$ then $q$ | ($\bar{p}$) | ($p$) | ($q$) | ($\bar{q}$) |
| | 81 | 29 | 75 | 40 |
| If not-$p$ then not-$q$ | ($\bar{p}$) | ($p$) | ($\bar{q}$) | ($q$) |
| | 77 | 27 | 46 | 63 |

[a] Correct choices, assuming full or defective implication.
*Key.* TA = true antecedent; FA = false antecedent; TC = true consequent; FC = false consequent.

In view of this parallel, I suggest that card selections do not reflect *any* process of reasoning, in the sense of analytic processing, and are due entirely to heuristic processes. In other words, subjects do not construct the problem space shown in Fig. 2 at all; they merely say that they would turn over the cards which appear, due to linguistic cues, to be 'relevant'. This hypothesis explains the logical inconsistency between performance on the selection task and that on other conditional reasoning tasks. The failure to select $\bar{q}$ conflicts with subjects' frequent recognition of TF as a falsifying case on the truth table task (cf. Table 2), and with their ability usually to draw correct *modus tollens* inferences of the form 'if $p$ then $q$, $\bar{q}$ therefore $\bar{p}$' (see Evans, 1982, chapter 8).

The assertion that card selections do not reflect any analytic processing also explains the puzzling results of 'therapy' experiments (e.g. Wason, 1969; Wason & Johnson-Laird, 1970). In these experiments subjects are induced to construct the problem space of Fig. 1

by the experimental procedure. They are asked, for example, to say what could be on the back of the $\bar{q}$ card, and what effect a $p$ would have on the truth-value of the rule. Subjects invariably say that a $p\bar{q}$ (TF) combination would falsify the rule, demonstrating the same analytic competence that the truth table task reveals. They are, however, notably reluctant to change their original selections despite substantial pressure from the experimenter and apparent self-contradiction (see Wason, 1977*a*). This strongly confirms the view that card selections are not based upon any such analysis of the consequences of turning the cards.

Another procedure which enforces analytic processing in the selection task is that of asking the subjects to provide verbal justifications of their choices. Discussion of this is deferred to a later section since it has implications for the 'dual process' theory of reasoning (Wason & Evans, 1975; Evans, 1982*a*, chapter 12).

Thus far, we have only considered performance on the abstract selection task. The well-known review by Wason & Johnson-Laird (1972) has led to a widespread belief that logical reasoning is facilitated on the selection task when the problem is phrased in thematic (concrete, realistic) terms. If this were indeed the case, I would have to suppose that selection on such problems were mediated by analytic processes. However, the two early studies which led to this conclusion (Wason & Shapiro, 1971; Johnson-Laird *et al.*, 1972) have not proved replicable in recent studies, and only certain types of thematic content reliably facilitate performance (see Griggs, 1983).

It now appears that simply concretizing the terms is not sufficient; the materials must directly stimulate the recollection of information relevant to the solution. For example, the widely quoted facilitation on the postal rule problem of Johnson-Laird *et al.* (1972) is not observed with American subjects who have never experienced a two-tier postal system (Griggs & Cox, 1982; Yachanin & Tweney, 1982) and with British subjects who are too young to remember it (Golding, 1981). Manktelow & Evans (1979) suggested that the original effect reflected retrieval of the correct answer from memory rather than a process of reasoning. Griggs & Cox (1982) provided positive confirmation of the 'memory-cue' hypothesis by achieving facilitation on a rule which was highly familiar to their subjects – the legal age rule for drinking alcohol in the State of Florida.

Thus, facilitation by thematic content can also be attributed to relevance judgements, though semantically rather than linguistically cued. You know *from experience* that the drinking laws are relevant if you are drinking alcohol or under the legal age, and irrelevant if you are drinking coke or over age. Like matching bias, memory cueing can be explained as a purely heuristic process.

In this section, the heuristic/analytic distinction has served several purposes. It has provided a basis for understanding the origins of logical and non-logical components of performance on the truth table task. It has explained both the similarities and the discrepancies between performance on this task and on the Wason selection task. In the latter case it has also helped to provide explanations of the abnormally poor logical performance observed on this task, the discrepancy between card selections and verbal evaluations, and the differential effects of various thematic contents.

*Statistical judgement*

The examples discussed so far involve logical reasoning. In order to show the generality of the heuristic/analytic distinction, I now consider some examples in a different sort of task: statistical judgement. The interest in this area lies in the extent to which people can make accurate intuitive judgements about probability and statistical information, without the help of formal procedures. There is much evidence of systematic biases (cf. Slovic *et al.*, 1977; Nisbett & Ross, 1980; Kahneman *et al.*, 1982). The dominant theoretical influence on this field over the past decade has been that of Kahneman & Tversky who propose that

such biases result from the use by subjects of heuristics such as 'representativeness' and 'availability'. There is evident risk of terminological confusion here since, as was stated earlier, my use of the term 'heuristic' differs markedly from theirs. In their approach heuristics are the basis of conscious strategies used to make decisions in place of application of normative rules of probability, which would be a type of analytic process.

In this section the focus is on certain phenomena which are amenable to the heuristic/analytic treatment in the sense that I use the terms. To anyone who has studied the literatures on both deductive reasoning and statistical judgement there are some striking parallels. As we have already seen, the picture subjects present of their logical abilities is highly variable across experimental settings. The same is true in statistical judgement studies. For example, do people understand the law of large numbers? Do they appreciate that larger samples provide more accurate information about populations than smaller ones? Some experiments suggest 'yes' and some 'no' (cf. Pollard & Evans, 1983).

The analogous explanation of this inconsistency to those offered for the discrepancies in logical reasoning experiments is as follows: subjects *do* understand the law of large numbers at the analytic level, but only apply this knowledge when heuristic processes determine that sample size is encoded as a relevant feature of the task information.

What factors might facilitate the (unconscious) selection process? Evans (1983*a*) suggests in the case of sample size that complexity of wording and perceptual salience are critical. In support of the former hypothesis it should be noted that certain problems, claimed by Kahneman & Tversky (1972*a*) to show subjects' lack of awareness of the role of sample size, produce normative performance when the wording of the problem is simplified (Evans & Dusoir, 1977; Bar-Hillel, 1979). Kahneman & Tversky (1972*a*) also claimed that with binomial samples subjects would concentrate on proportionality to the exclusion of sample size. In fact, when given simple paired comparisons between samples in which both factors are equally salient, a substantial minority gave more weight to the evidence of larger samples (Evans & Dusoir, 1977; Evans & Pollard, 1982).

Investigation of this problem has also brought out the importance of perceptual factors in the selective processing of task information. For example, Olson (1976) showed that subjects asked to generate sampling distributions attended neither to sample size *nor* to sample proportion. They instead produced a characteristic bell-shaped distribution over the available response categories, regardless of what the categories actually represented. In an unpublished study of judgements about normally distributed data, Evans & Pollard (described by Evans, 1983*a*) found that subjects attended to sample size but ignored sample variability when the data were displayed as blocks of numbers, with reverse trends when the same data were presented as frequency histograms. Whilst the two forms of display enhanced the salience of one or other feature of the sample, the information about both was available in either. This suggests that subjects are rather passive in their approach to inferential tasks. They apply analytic procedures to task features which appear relevant, but they do not actively seek relevant data.

The salience of presented features, while important, is clearly not *sufficient* to determine relevance. In the Wason selection task, for example, all four cards are salient, but two are generally ignored as irrelevant, unless favoured by linguistic or semantic cues. A comparable problem in statistical judgement, where semantic factors are critical to the perceived relevance of a feature, arises in the base-rate problem. This problem has generated considerable interest and debate in the recent literature (e.g. Cohen, 1981, and associated commentaries). The problem is the essentially Bayesian one that any specific statistical evidence about a hypothesis must be weighted by consideration of prior probabilities or base-rates.

Kahneman & Tversky (1973) presented several experiments in which subjects ignored

base-rate data. In some of these the relevant data were made highly salient. For example, subjects' estimates of the likelihood of someone being a lawyer or an engineer on the basis of a personality description were unaffected by accompanying information that he was drawn from a group composed mostly of engineers (or lawyers). This was true if *any* information was provided about the individual, even if totally non-diagnostic.

I suggest that the reason for this is that people (and indeed some philosophers) do not see frequency data as relevant to the probability of an *individual* case. Certainly, people seem remarkably unimpressed by actuarial data about the effects of cigarette smoking, non-wearing of seat belts, etc. The 'it can't happen to me' belief may have a cognitive, rather than motivational, base. Many people seem to accept that smokers *as a group* are more prone to disease, while believing that one cannot infer anything about an individual such as themselves. The root cause is the fundamental problem with the concept of probability; no *particular* thing happens with a probability of 0·7: it either happens or it does not.

There have been a number of studies of the so-called base-rate fallacy (for reviews see Nisbett & Ross, 1980, chapter 7; and Pollard & Evans, 1983). Whilst the fallacy has been replicated a number of times, our interest lies in the cases where subjects *do* take account of base-rates. As with the Wason selection task it is the semantic context which is crucial to 'facilitation'. Ajzen (1977) and Tversky & Kahneman (1980) have shown that a major factor is the perception of a *causal* basis for the effect of the base-rate. For example, in the well known 'cabs' problem subjects normally estimate the chance that a witness has correctly identified a cab's colour to be equal to the stated reliability of the witness, regardless of information about the number of blue and green cabs in the city (Kahneman & Tversky, 1972*b*). If, however, subjects are told that, while the number of cabs of either colour are the same, far more green (or blue) ones are involved in accidents, they shift their judgements in accordance with the base-rate data (Tversky & Kahneman, 1980). While statistically equivalent, there is a crucial semantic difference. In the latter case there is a causal link: green cab drivers are evidently reckless and cause more accidents to happen.

Bar-Hillel (1980) has suggested that it is not causality *per se* but perceived *relevance* of base-rate data that is important. She suggests that non-causal base data will influence subjects if they are made to appear relevant in other ways. Whilst this argument is evidently conducive to my own view, there are problems with her particular experiments (see Pollard & Evans, 1983).

These examples of statistical judgement research produce close parallels to the deductive reasoning examples discussed earlier. It appears that any analytic judgemental or reasoning process is based only on features of the task content which have been encoded as 'relevant' to the task. The pre-attentive, heuristic processes responsible for this selection are themselves influenced by a number of factors, including perceptual salience, linguistic and semantic cues. We now turn to a broader consideration of the theoretical significance of this analysis.

## Theoretical implications

### *Relation to the dual process theory*

Wason & Evans (1975) reported an experiment and presented a theory which grew from it and a lot of other relevant work in the literature. The experiment was designed to reconcile some apparently conflicting findings of earlier papers. Johnson-Laird & Wason (1970) proposed that responses on the Wason selection task reflect the degree of 'insight' held by the subject. For example, a subject with partial insight correctly seeks falsifying combinations but also seeks, unnecessarily, to verify the rule. This results in a choice of the cards $p$, $q$ and $\bar{q}$. Other levels of insight lead to other choices. In order to provide a

non-circular test of the theory, Goodwin & Wason (1972) inferred the level of insight from verbal justifications provided by the subjects. These correlated with selections as predicted by the model.

The Johnson-Laird & Wason (1970) model does, however, assume Wason's original explanation of the selection pattern *p* and *q* as being due to verification bias. As we saw in an earlier section this theory was apparently undermined by Evans & Lynch's (1973) demonstrations of matching bias. In order to resolve the conflict, Wason & Evans (1975) combined the methodology of both studies; subjects solved a problem with a rule of the form 'if *p* then *q*' *and* one of the form 'if *p* then not-*q*', in each case providing verbal explanations for their choices. In accordance with matching most subjects gave the wrong solution on the affirmative rule, but the correct solution (*p* and *q*) on the negative rule. The verbalizations showed a high degree of insight on the negative rules – i.e. a desire to falsify the rules – but no insight (verification bias) on the affirmative rules, even when presented as the second task.

Wason & Evans (1975) concluded that the verbalizations did not reflect insight into the actual choices, but were *post hoc* rationalizations (see also Evans & Wason, 1976). It was argued that there were two distinct types of thought involved: *type 1* which underlay decisions and *type 2* which underlay verbal justifications. Wason & Evans discussed a number of previous studies which also suggested the operation of dual processes.

Now, as the reader may suspect, there is a close correspondence between the type 1 and type 2 processes of Wason & Evans, and the 'heuristic' and 'analytic' processes of the present account. There are, however, a number of unsatisfactory features of the original theory (and its development by Evans, 1982*a*, chapter 12) which can now be resolved.

In the first place, Wason & Evans assumed that justifications reflected entirely type 2 processes. The latter assumption is hard to justify in the light of one aspect of their data, which they called 'secondary matching bias'. In the protocols, subjects tended to refer to the possible matching value on the *other* side of the card. If there are dual processes, surely they will not both be subject to matching bias? I now propose that the task demands of verbal justification produce additional heuristic as well as analytic processing.

In the section on the selection task, it was proposed that card selections were *purely* heuristic: all the subject did was to choose cards which appeared relevant. Thus subjects did not look beyond the first level of the problem space (Fig. 2). When asked to provide verbal justifications, however, subjects *then* generated the next level, i.e. possible values on the backs of the cards. This is subject to the same heuristic processes as before, leading to a focus on matching values on the other side of the card (secondary matching bias). The evaluation of card combinations is then provided by the analytic system, which assigns truth-values in a manner consistent with the true/false evaluations on the truth table evaluation task. For example, subjects may be asked to justify the choice of *p* on the rule 'if *p* then *q*'. They will typically say that a *q* on the back would prove the rule true. However, on justifying the choice *p* on the rule 'if *p* then not-*q*', they will typically say that a *q* on the back would make the rule false. In each case they think about the possibility of *q* rather than $\bar{q}$ – due to heuristic selection. The explanation in terms of verification or falsification accords with a logical analysis of the combination produced in each case.

The 'rationalization' referred to by Wason & Evans is then the application of the analytic system to show the logical consequences of decisions *already taken*. Some comparable data on a different task are provided by Evans *et al.* (1983). They analysed both thinking aloud and retrospective protocols on a syllogistic reasoning task, where the conclusions to arguments had clear empirical truth-value. Subjects who showed a strong belief bias, i.e. claimed 'true' conclusions to be valid and 'false' conclusions to be invalid, regardless of the premises, showed characteristic features in their protocols. When logic

accorded with belief, they referred to the premises presented, but when it conflicted with belief they referred instead to extraneous information. The fascinating aspect of this is that, since their protocols discriminate the validity of their problems, they must in some way have understood the logic. For example, if you believe the conclusion to be true and declare it 'valid', you only then justify with reference to the premises *if* they support the conclusion logically. As in the Wason & Evans (1975) experiment, subjects are showing a greater degree of logical (analytic) ability in their protocols, than in their actual choices. Presumably belief bias, like matching, operates through the heuristic system.

Wason & Evans' assumption that card selections were entirely due to type 1 processes corresponds to the assertion of the present paper that they are entirely heuristic. However, this analysis *conflicts* with the claim of Evans (1977) that selections reflect parallel, competing 'logical' and 'non-logical' components, and with the revised dual process theory presented by Evans (1982*a*, chapter 12). The parallel components in the Evans (1977) stochastic model are the tendencies to match and the tendency to prefer true over false antecedent selections. The latter was thought by Evans (1977, 1982*a*) to reflect the logical component of performance. However, in the present paper this has been assumed to be a second linguistic bias, in addition to matching, which acts to determine perceived relevance.

It is now suggested that the original form of the dual process theory in which type 1 (heuristic) processes are followed by type 2 (analytic) processes is preferable to the revised theory presented by Evans (1982*a*). The Evans (1977) model is seen *not* as an account of how 'logical' and 'non-logical' tendencies combine, but rather as a description of the parallel determination of heuristic card selections by the cues of linguistic relevance. Thus, while parallel processes may operate *within* the heuristic stage, there is nevertheless a sequential (or alternating) relationship between heuristic and analytic stages of processing. It also seems improbable that parallel processing would operate at the analytic stage. The heuristic/analytic theory, while encompassing the proposals of Wason & Evans, goes well beyond them. While preconscious heuristic processes have a large influence on inferential behaviour, we have also seen that analytic processes bear on the actual decision taken in many cases.

*Analytic processes and the rationality debate*

The discussion so far has been intended to demonstrate the existence of heuristic and analytic stages, and to illustrate some of the factors which can determine heuristic selection of task features for further analytic processing. We must now consider the possible nature of these analytic processes, not only for theoretical completeness, but also to relate the theory to the current debate about human rationality.

The idea that behaviour may be irrational seems to derive from the notion of *bias* in studies of human inferences. A bias is a source of error which is systematic rather than random, and consists of either failure to take account of a normatively relevant feature or else a tendency to respond to a normatively irrelevant feature. Note that this definition describes behaviour in relation to some *a priori* norm, a procedure which Cohen (1982) calls the 'preconceived norm method'.

One of the principal critics of the practice of describing subjects' behaviour as irrational and biased is, in fact, Cohen (1979, 1981, 1982). The two major components of his argument are that (1) one must distinguish between underlying competence and observed performance and (2) the preconceived normative rule of inference of the experimenter may not be that of the subject; hence subjects may be rational relative to their own system of rules, whilst being described as irrational by the experimenter.

My own view is that the task of psychologists is to understand the nature of inferential behaviour rather than to judge it as rational or irrational (cf. Evans & Pollard, 1981;

Griggs, 1981). However, I do believe that a motivation to attribute rationality to subjects has often led psychologists to provide inaccurate descriptions of such behaviour (numerous examples are discussed by Evans, 1982*a*). For example, following Henle (1962) some psychologists have sought to explain illogical performance on the assumption that subjects reason logically but with a misinterpretation of the problem information – an approach which provides a very incomplete description of subjects' behaviour.

Now, the heuristic/analytic distinction proposed here might be seen as a device for preserving the notion of logical competence. The argument would be that the analytic procedures involve the application of rule systems to data encoded as relevant. In line with Cohen's arguments, 'errors' would reflect either performance factors operating through the heuristic system, or else the adoption of non-standard rule systems in the analytic system. This view might also be reconcilable with those of others who believe that reasoning is achieved by the application of a general set of inferential rules (e.g. Braine, 1978; Rips, 1983*a*, *b*).

This is *not*, however, what I wish to propose. Like Wason (1977*b*) and Johnson-Laird (1983), I do not believe that reasoning is generally achieved by the application of 'natural' inference rules to underlying abstract propositional structures (though, evidently, philosophers and others may be trained to apply such rules to formally stated problems). I also agree with others (e.g. Pollard, 1982; Griggs, 1983; Wason, 1983) who believe reasoning processes to be highly content and context dependent. Thus, while analytic processes serve to generate inferences, they do not normally do so by a process of logical reasoning. To see why, we must re-examine some of the phenomena described earlier in the paper.

In the truth table task, the 'competence' exhibited consists of a tendency to classify TT as conforming to the rule and TF as falsifying it among those cases not discarded as irrelevant (cf. Table 2). There is also a tendency to classify FT as falsifying, which is stronger on some rules than others. Performance on conditional inference tasks shows some consistency with this (cf. Evans, 1982*a*, chapter 8). For example, people frequently endorse *modus ponens* and *modus tollens* inferences, though the latter is susceptible to a response bias, and have a weaker tendency to endorse those inferences which assume FT to describe a proscribed situation.

These 'analytic' decisions, while consistent with a form of logical analysis, need not reflect an internalized set of abstract logical rules. The trends referred to above occur with artificial experimental materials such as letters and numbers. When presented with such materials, it is proposed, subjects attempt to draw inferences by consideration of the *normal linguistic usage* of the syntactic form, in this case the 'if–then' rule. To see that you must have a *q* with a *p* is no more than the application of knowledge as to what the 'if–then' sentence usually means in natural language (see also Pollard & Evans, 1980, but for some counter-examples see Johnson-Laird, 1983, chapter 3). In line with this linguistic usage hypothesis, I have argued (Evans, 1982*a*, chapters 7 and 8) that the 'if–then' sentence is ambiguous with abstract content because in natural language *context* determines either an implication or equivalence reading. For an excellent illustration of the manner in which context influences inferences with thematic conditional rules, see Fillenbaum (1975, 1976).

What I am saying, in the case of verbal reasoning, is that analytic processes reflect the operation of linguistic and semantic systems, rather than an abstract logic. When problems are presented in thematic terms which can be related to the subjects' experience, inferences are generated by direct or analogical use of that experience. Johnson-Laird & Wason (1977, pp. 341–354) distinguish between 'conscious deductions' and the sorts of implicit inference which are involved in understanding sentences. I would argue that this is a false

distinction: the processes of verbal reasoning are inextricably linked with those of understanding sentences in relation to real-world knowledge. The fact that the *task set* demands an explicit inference does not mean that a different type of process is elicited. Where materials are either abstract or arbitrarily thematic, however, clearly some form of analogical reasoning is involved based on the linguistic structure. Subjects ask (not necessarily *consciously*) in what context would this type of sentence normally be used, and what kinds of inference would normally be appropriate in such a context?

Verbal inferences, then, reflect subjects' *knowledge* of language and of particular contexts; they reason from experience rather than from logic – assuming, that is, that analytic processes are engaged. It is worth noting that, while I agree, in general, with the idea of contextually bound reasoning proposed by Wason (1983), I disagree with his application of this idea to explain performance on the thematic selection task. As was argued in an earlier section, performance on that particular task is proposed to be entirely heuristic.

What, however, of non-verbal reasoning, such as statistical inference? For example, Kahneman & Tversky (1982*a*) are somewhat ambivalent about subjects' understanding of the role of sample size:

> It has been demonstrated that adults do not have generally valid intuitions corresponding to the law of large numbers... But it is simply not the case that every problem to which these rules are relevant will be answered incorrectly, or that the rules cannot appear compelling in certain contexts.

The inconsistency which troubles them could be resolved in two different ways. On the one hand it may be that subjects *do* possess such a rule, but only apply it when heuristic processes have selected sample size as a relevant task feature. Alternatively, if analytic processes are based on applications of specific knowledge, rather than general rules, then they may generate a 'correct' use of sample size in some contexts, but not others.

My own view is a combination of these. Although the relevance of sample size is determined heuristically, the analysis of its effects on the inference is determined by applications of specific knowledge. However, the knowledge used may be either explicitly verbal or intuitive. In the case of sample size, subjects may recall the law of large numbers from statistics classes as an explicit verbal rule and apply it. Alternatively, they may have relevant experience which leads them 'intuitively' to prefer larger samples. In either case, however, the inferential process would be context-related. Students of logic and statistics notoriously fail to apply the rules they have been taught when the problem is disguised in an unfamiliar form. Similarly, intuitive convictions are likely to be limited to contexts where the relevance of previous experience is recognizable.

One reason why it does not appear that subjects have a reliable general rule about sample size is that they can be induced to make *inappropriate* use of this feature. Bar-Hillel (1979), for example, found that subjects given information about sample *and* population size based judgements on sample to population ratio, rather than the normatively prescribed absolute sample size. A problem here is that the mere presentation of information in a psychological experiment is a cue to 'relevance'. The subject is likely to think that if population size is presented it must be relevant to the task and try to construct inferences which make use of it. Cohen (1982) makes a plausible suggestion as to how subjects might do this.

It is now clear that the terms 'reasoning' and 'analytic processing', as I use them, refer to *any* process used to derive inferences or judgements from 'relevant' data, and do not presuppose the possession of inference rules or 'mental logics' (cf. Johnson-Laird, 1983). Indeed, some of the processes which Tversky & Kahneman describe as 'heuristic' would be analytic in my terms. For example, if subjects *do* judge probabilities in some situations by

the ease with which they can bring examples to mind (the 'availability heuristic'), I would have to call that analytic. People are here making use of their past experience and knowledge to generate a judgement. Similarly, the 'simulation heuristic' proposed by Kahneman & Tversky (1982*c*) suggests a plausible method that subjects might use to *analyse* the likelihood of certain events, by constructing scenarios. This means, as Evans (1982*b*) implies, that Kahneman & Tversky's (1982*a*) distinction between 'rules' and 'heuristics' is a false one. The distinction is based on the (preconceived) normative appropriateness of the behaviour produced rather than on the nature of the psychological process involved.

Where, then, does this leave the issue of 'rationality'? If rational reasoning entails the possession and application of formal logical rules, then I do not believe people to be rational, even in their analytic processing. On the other hand, if rational reasoning is the application of such knowledge and analogous experience as one can muster to assess the probable inferences that can be drawn, then people *are* rational. This rationality is, however, 'bounded' in two distinct ways. Cognitive constraints limit both what features are selected as relevant, and the appropriateness of the inferences derived from thinking about those features.

*Relation to other theories of inference*

A popular contemporary approach to reasoning theory involves the construction and empirical test of formal models which are applied to specific experimental paradigms (e.g. Sternberg, 1980; Guyote & Sternberg, 1981). Such models are very precise but suffer an evident risk of tunnel vision. In this context, Sternberg's (1982) recent broadening of perspective for his models is most welcome.

The heuristic/analytic framework bears a metatheoretical relationship to such specific models. In formal terms, if a set of features are potentially available to the subject, then the weighting attached to each feature would be determined by heuristic processes, while the precise function of a given cue would depend upon the analytic system. The problem with formal modelling, however, is that it may be premature and focus too much attention on specific tasks. Many of the theoretical distinctions drawn in the empirical review section of this paper required a broad view of a range of experimental situations.

Turning to more general theoretical approaches, there is firstly the 'heuristics and biases' approach to judgement, largely popularized by Kahneman and Tversky. Several examples of this approach have already been discussed, and the difference between their 'heuristics' and mine explained. Pollard (1982) has presented an argument that their 'availability' heuristic can be applied to the explanation of deductive reasoning data. His use of the term 'availability' differs from the concept of 'relevance' in the present paper. As was implied earlier, availability of features, though necessary, is not sufficient for the subject to code them as relevant. Also, Pollard uses availability as a direct explanation of reasoning responses and does not pose a second (analytic) stage or, indeed, any process of reasoning as such. It is interesting to note, however, that in a recent paper Kahneman & Tversky (1982*a*) have proposed a distinction not unlike the one expounded in this paper. They distinguish between errors of *comprehension*, in which the subjects do not understand the correct rule, and errors of *application* where they understand but do not use the rules. If 'rule application' is taken to be a form of analytic processing, then my heuristic processes would determine the likelihood of their application.

I do, however, believe that Kahneman & Tversky placed too much emphasis on the concept of normative rules in this account (cf. Evans, 1982*b*), and have already rejected the notion that analytic reasoning is generally based upon natural inference rules (as, for example, in Braine, 1978; and Rips, 1983*a*, *b*). In this latter view I am in total accord with

Johnson-Laird (1983) who expounds an ingenious account of how people reason without logic.

Johnson-Laird proposes that people construct 'mental models' to represent possible situations described by verbal propositions. In an explicit deduction problem, if they can construct no models consistent with the premisses in which a proposition would be false, then they may infer that proposition to be a valid inference. Johnson-Laird also proposes a number of psychological constraints on deductive abilities. For example, working memory capacity limits subjects' ability to keep track of more numerous or more complex models. Also the manner in which models are constructed may lead to biases in the processing of the information they contain.

Johnson-Laird's theory has no serious incompatibility with the present framework. Not only does he reject inference rules, but his theory of deduction (one type of analytic process) is essentially semantic, and thus well able to cope with content-specific reasoning. His emphasis on the importance of problem representation (via mental models) is also compatible with the present stress on heuristic processes. The selection of items as 'relevant' is a concept closely related to that of problem representation. While specific applications of his theory to reasoning data have so far been concentrated mostly on categorical syllogism tasks, this is not an inherent limitation in his approach. For example, Kahneman & Tversky's (1982*b*) 'simulation heuristic', which involves mental simulations, is clearly compatible with the theory of mental models, even though its output is a probabilistic rather than deductive inference.

## Conclusions

The main purpose of this paper was an attempt to assist future research on the topic of human inference by drawing a simple, but important, distinction between 'heuristic' and 'analytic' processes. While these terms themselves are less than ideal (I could find none better) the concepts they represent are clear. Both in the discussion of experimental research and alternative theories it has been argued that misinterpretations and confusions have arisen from a failure to distinguish clearly between *how* people reason, and *what* they reason about. The recent debate about rationality has, if anything, obscured the real issues in the field.

How, then, can the present framework help to generate productive study of human inference? Firstly, we must recognize that there is little point in assessing inferences unless we know what information they are based upon. Experiments must be designed to isolate the effects of various features upon inferential behaviour. It is not only theoretically interesting to understand the nature of heuristic selection, but also practically important. For example, how can problems be presented to decision makers so as to maximize their attention to relevant data and minimize their concern with irrelevant data? Once this is understood, then the manner in which psychologically 'relevant' information is processed can be studied. Only then can we argue sensibly about whether this is achieved by inference rules, mental models or whatever. Whilst my own present preference lies with exploration of the heuristic system, those who prefer to study the reasoning mechanisms themselves should first ensure that they know wherein the focus of their subjects' attention lies.

## Acknowledgements

The writing of this paper benefited from discussions with Richard Griggs and Paul Pollard whose contributions are gratefully acknowledged.

## References

Ajzen, I. (1977). Intuitive theories of events and the effects of base rate information on prediction. *Journal of Personality and Social Psychology*, **35**, 303–314.

Bar-Hillel, M. (1979). The role of sample size in sample evaluation. *Organisational Behavior and Human Performance*, **24**, 245–257.
Bar-Hillel, M. (1980). The base rate fallacy in probability judgements. *Acta Psychologica*, **44**, 211–233.
Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, **85**, 1–21.
Cohen, L. J. (1979). On the psychology of prediction: Whose is the fallacy? *Cognition*, **7**, 385–407.
Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioral and Brain Sciences*, **4**, 317–370.
Cohen, L. J. (1982). Are people programmed to commit fallacies? Further thought about the interpretation of experimental data on judgement. *Journal for the Theory of Social Behavior*, **12**, 251–274.
Einhorn, H. J. & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, **32**, 53–88.
Evans, J. St B. T. (1972*a*). On the problems of interpreting reasoning data: Logical and psychological approaches. *Cognition*, **1**, 373–384.
Evans, J. St B. T. (1972*b*). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, **24**, 193–199.
Evans, J. St B. T. (1975). On interpreting reasoning data: A reply to Van Duyne. *Cognition*, **3**, 387–390.
Evans, J. St B. T. (1977). Toward a statistical theory of reasoning. *Quarterly Journal of Experimental Psychology*, **29**, 621–635.
Evans, J. St B. T. (1982*a*). *The Psychology of Deductive Reasoning*. London: Routledge & Kegan Paul.
Evans, J. St B. T. (1982*b*). On statistical intuitions and inferential rules: A discussion of Kahneman & Tversky. *Cognition*, **12**, 319–324.
Evans, J. St B. T. (1983*a*). Selective processes in reasoning. In J. St B. T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan Paul (in press).
Evans, J. St B. T. (1983*b*). Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology*, **35A**, 635–644.
Evans, J. St B. T., Barston, J. & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, **11**, 295–306.
Evans, J. St B. T. & Dusoir, A. E. (1977). Proportionality and sample size as factors in intuitive statistical judgement. *Acta Psychologica*, **41**, 129–137.
Evans, J. St B. T. & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology*, **64**, 391–397.
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. & Pollard, P. (1981). On defining rationality unreasonably. *The Behavioral and Brain Sciences*, **4**, 35–36.
Evans, J. St B. T. & Pollard, P. (1982). Statistical judgement: A further test of the representativeness construct. *Acta Psychologica*, **51**, 91–103.
Evans, J. St B. T. & Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology*, **67**, 486–497.
Fillenbaum, S. (1975). If: Some uses. *Psychological Research*, **37**, 245–260.
Fillenbaum, S. (1976). Inducements: On phrasing and logic of conditional promises, threats and warnings. *Psychological Research*, **38**, 231–250.
Golding, E. (1981). The effect of past experience on problem solving. Paper read to The British Psychological Society at Surrey University.
Goodwin, R. Q. & Wason, P. C. (1972). Degrees of insight. *British Journal of Psychology*, **63**, 205–212.
Griggs, R. A. (1981). Human reasoning: Can we judge before we can understand? *The Behavioral and Brain Sciences*, **4**, 38–39.
Griggs, R. A. (1983). The role of problem content in the selection task and THOG problem. In J. St B. T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan Paul.
Griggs, R. A. & Cox, J. R. (1982). The elusive thematic materials effect in Wason's selection task. *British Journal of Psychology*, **73**, 407–420.
Guyote, M. J. & Sternberg, R. J. (1981). A transitive chain theory of syllogistic reasoning. *Cognitive Psychology*, **13**, 461–525.
Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, **69**, 366–378.
Johnson-Laird, P. N. & Wason, P. C. (1977). *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.
Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge: Cambridge University Press.
Johnson-Laird, P. N., Legrenzi, P. & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, **63**, 395–400.
Johnson-Laird, P. N. & Tagart, J. (1969). How implication is understood. *American Journal of Psychology*, **2**, 367–373.
Johnson-Laird, P. N. & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, **10**, 64–98.
Kahneman, D., Slovic, P. & Tversky, A. (eds) (1982). *Judgement under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Kahneman, D. & Tversky, A. (1972*a*). Subjective probability: A judgement of representativeness. *Cognitive Psychology*, **3**, 430–454.
Kahneman, D. & Tversky, A. (1972*b*). On prediction and judgement. *ORI Research Monograph*, **12** (4).
Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, **80**, 237–251.
Kahneman, D. & Tversky, A. (1982*a*). On the study of statistical intuition. *Cognition*, **11**, 123–142.
Kahneman, D. & Tversky, A. (1982*b*). A reply to Evans. *Cognition*, **12**, 325–326.
Kahneman, D. & Tversky, A. (1982*c*). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (eds), *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
Legrenzi, P. (1970). Relations between language and reasoning about deductive rules. In G. B. Flores D'Arcais & W. J. M. Levelt (eds), *Advances in Psycholinguistics*. Amsterdam: North Holland.
Manktelow, K. I. (1982). Recent developments in research on Wason's selection task. *Current Psychological Reviews*, **1**, 257–268.
Manktelow, K. I. & Evans, J. St B. T. (1979). Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology*, **70**, 477–488.
Marcus, S. L. & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior*, **18**, 199–223.
Newell, A. (1980). Reasoning, problem solving and decision processes: The problem space as a fundamental category. In R. S. Nicherson (ed.), *Attention and Performance*, VIII. Hillsdale, NJ: Erlbaum.
Newell, A. & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
Nisbett, R. E. & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgement*. Englewood Cliffs, NJ: Prentice-Hall.
Olson, D. R. (1976). Some apparent violations of the representativeness heuristic in human judgement. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 599–608.
Pollard, P. (1982). Human reasoning: Some possible effects of availability. *Cognition*, **12**, 65–96.
Pollard, P. & Evans, J. St B. T. (1980). The influence of logic on conditional reasoning performance. *Quarterly Journal of Experimental Psychology*, **32**, 605–624.
Pollard, P. & Evans, J. St B. T. (1983). The role of 'representativeness' in statistical inference: A critical appraisal. In J. St B. T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan Paul.
Rips, L. J. (1983*a*). Reasoning as a central intellective ability. In R. J. Sternberg (ed.), *Advances in the Study of Human Intelligence*, vol. 2. Hillsdale, NJ: Erlbaum.
Rips, L. J. (1983*b*). Cognitive processes in propositional reasoning. *Psychological Review*, **90**, 38–71.
Rips, L. J. & Marcus, S. L. (1977). Suppositions and the analysis of conditional sentences. In M. A. Just & P. A. Carpenter (eds), *Cognitive Processes in Comprehension*. New York: Wiley.
Slovic, P., Fischhoff, B. & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, **28**, 1–39.
Sternberg, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General*, **109**, 119–159.
Sternberg, R. J. (1982). Reasoning, problem solving and intelligence. In R. J. Sternberg (ed.), *Handbook of Human Intelligence*. Cambridge: Cambridge University Press.
Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, **185**, 1124–1131.
Tversky, A. & Kahneman, D. (1980). Causal schemata in judgements under uncertainty. In M. Fishbein (ed.), *Progress in Social Psychology*. Hillsdale, NJ: Erlbaum.
Wason, P. C. (1966). Reasoning. In B. N. Foss (ed.), *New Horizons in Psychology I*. London: Penguin.
Wason, P. C. (1969). Regression in reasoning. *British Journal of Psychology*, **60**, 471–480.
Wason, P. C. (1977*a*). Self-contradictions. In P. N. Johnson-Laird & P. C. Wason (eds), *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.
Wason, P. C. (1977*b*). The theory of formal operations: A critique. In B. Geber (ed.), *Piaget and Knowing*. London: Routledge & Kegan Paul.
Wason, P. C. (1983). Realism and rationality in the selection task. In J. St B. T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan Paul.
Wason, P. C. & Evans, J. St B. T. (1975). Dual processes in reasoning? *Cognition*, **3**, 141–154.
Wason, P. C. & Johnson-Laird, P. N. (1970). A conflict between selecting and evaluating information in an inferential task. *British Journal of Psychology*, **61**, 509–515.
Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content*. London: Batsford.
Wason, P. C. & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, **23**, 63–71.
Yachanin, S. A. & Tweney, R. D. (1982). The effect of thematic content on cognitive strategies in the four-card selection task. *Bulletin of the Psychonomic Society*, **19**, 87–90.

*Received 3 August 1983; revised version received 21 October 1983*

Requests for reprints should be addressed to Jonathan St B. T. Evans, Department of Psychology, Plymouth Polytechnic, Plymouth PL4 8AA, UK.