Manuscript Information	
Journal Acronym	PTAR
Volume and issue	
Author name	
Manuscript No. (if applicable)	_A_241481

Thinking & Reasoning

Typeset by
KnowledgeWorks Global Ltd.
for
Psychology Press
Taylor & Francis Group

KNOWLEDGEWORKS GLOBAL LTD.

QUERIES: to be answered by AUTHOR

Dear Author

Please address all the numbered queries on this page which are clearly identified on the proof for your convenience.

Thank you for your cooperation

QUERY NO.	QUERY DETAILS
1	Not listed in References.

THINKING & REASONING, 0000, 00 (0) 000–000

Psychology Press
Taylor & Francis Group

Conditional reasoning and the Wason selection task: Biconditional interpretation instead of reasoning bias

Pascal Wagner-Egger
University of Fribourg, Switzerland

Two experiments were conducted to show that the IF ... THEN ... rules used in the different versions of Wason's (1966) selection task are not psychologically—though they are logically—equivalent. Some of these rules are considered by the participants as strict logical conditionals, whereas others are interpreted as expressing a biconditional relationship. A deductive task was used jointly with the selection task to show that the original abstract rule is quite ambiguous in this respect, contrary to deontic rules: the typical "error" made by most people may indeed be explained by the fact that they consider the abstract rule as a biconditional. Thus, there is no proper error or bias in the selection task as it is still argued, but a differential interpretation of the rule. The need for taking into account a pragmatic component in the process of reasoning is illustrated by the experiments.

For a long time, the meaning of the expression IF ... THEN ... has been a major focus in reasoning and psycholinguistic research (e.g., Braine & O'Brien, 1998; Fillenbaum 1975; Geis & Zwicky, 1971; Johnson-Laird & Byrne, 2002; Legrenzi, 1970; Oaksford & Chater, 2003). In natural language, the expression IF ... THEN ... can either refer to a conditional relationship (e.g., *If it is a canary, then it is yellow*), or to a biconditional one (e.g., *If it is a dog, then it barks*), depending on the content of the propositions (Wason & Johnson-Laird, 1972). The cognitive psychology of

Correspondence should be address to Pascal Wagner-Egger, Department of Psychology, University of Fribourg, Rue de Faucigny 2, 1700 Fribourg, Switzerland.
E-mail: pascal.wagner@unifr.ch

I would like to thank Pascal Gygax and Adrian Bangerter for their suggestions on the manuscript. I am also indebted to Richard Griggs, Vittorio Girotto, Ken Manktelow, and an anonymous reviewer for their helpful comments on earlier versions of this paper.

© 2007 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/tar DOI: 10.1080/13546780701415979

reasoning aims at studying how lay persons carry out deductions, in particular those from conditional rules, such as:

1. *If p then q*, $p \Rightarrow q$ (Modus Ponens = MP)
2. *If p then q*, *not-p* $\Rightarrow$ *not-q* (fallacy of Denying the Antecedent = DA)
3. *If p then q*, $q \Rightarrow p$ (fallacy of Affirming the Consequent = AC)
4. *If p then q*, *not-q* $\Rightarrow$ *not-p* (Modus Tollens = MT)

Whereas deductions (1) and (4) are logically valid, (2) and (3) are not. But if the first premise is a biconditional rule *if and only if p then q*, all four deductions are valid (as can be seen by trying them out with "the canary" and "the dog").

Despite a large variation across experiments, it is widely accepted that abstract IF ... THEN ... rules are often treated as biconditionals (Evans, Newstead, & Byrne, 1993). Wason (1964, in Wason & Johnson-Laird, 1972) suggested that the two fallacies (DA and AC) were due to illicit conversions or inversions of the rule *if p then q* to *if q then p* for the fallacy of affirming the consequent, and to *if not-p then not-q* for the fallacy of denying the antecedent. These two conversions correspond logically to a biconditional reading of the IF ... THEN ... rule.

THE CONVERSION OF CONDITIONALS

The phenomenon of a biconditional interpretation of IF ... THEN ... expressions has been widely documented in reasoning literature concerning syllogistic reasoning research (Wilkins, 1928), Piagetian research (Matalon, 1962), or developmental psychology (Girotto, 1990). A crucial question regarding the conversion of the conditional rule is whether it has to be considered as a reasoning error. According to Polya (1954, in Wason & Johnson-Laird, 1972), though not valid from the standpoint of formal logic, such conversions should be treated as "plausible" instead of erroneous. Chapman and Chapman (1959) also suggested that these inferences were to be considered as normal, for the simple reason that numerous examples in language render them possible. Moreover, people transfer them into abstract rules in laboratory settings. Geis and Zwicky (1971) introduced the concept of "invited inferences", or even "conditional perfection", to name these so-called fallacies. Fillenbaum (1975) also advocated that pragmatic inferences rooted in social context might alter the logical structure of propositions. In particular, he drew attention to the fact that IF ... THEN ... expressions have numerous interpretations depending on their content (e.g., promise, threat, temporal relationship, etc.). In essence, the expression IF ... THEN ... is a true "chameleon" whose colour changes according to its environment.

In light of these arguments, it seems unwise to consider these context- and content-dependent inferences as fallacies or biases from the point of view of formal logic, which is precisely blind to content and context. Besides the conditional reasoning task described above, one of the most famous experimental paradigms that deals with the conditional rule is undoubtedly Wason's selection task (Wason, 1966, 1968).

THE WASON SELECTION TASK

In its abstract version, four cards are presented to participants, on which are respectively two letters, for example the letters E and K, and two numbers, for example the numbers 4 and 7. After being informed that each card has a letter on one side and a number on the other, the participants have to indicate which card they would turn over in order to see whether the conditional rule *If there is a vowel on one side, there is an even number on the other side* is true or false. Most participants choose cards E and 4. In terms of logic, these choices represent cases *p* & *q* if the rule is stated *if p then q*, which is considered to be an error from the viewpoint of formal logic, whereas the correct response would be to select cards E and 7, *p* & *not-q*.

Despite its apparent simplicity, this problem has proved to be difficult for participants. Indeed, only 4% reached the correct response in the first experiments summarised by Wason and Johnson-Laird (1972): 33% of the participants selected the response *p* alone, and 46% the response *p* & *q*. Nonetheless, the task itself has also been widely challenged by researchers. Several researchers have suggested that the dominant selection of the *p* & *q* cards could be considered as correct under a probabilistic analysis of the problem as decision making (e.g., Evans & Over, 2004; Oaksford & Chater, 1994), thus questioning the appropriateness of formal logic as a normative standard.

Consequently, different theoretical analyses of the task have been proposed. First, the *p* & *q* selection has been considered as a result of a *hypothesis confirmation bias* (Wason & Johnson-Laird, 1972). In essence, the participants would prefer to confirm than to refute their hypotheses, as non- or pseudo-scientists do (e.g., Popper, 1959). But as some of Popper's critics have largely documented (Kuhn, 1962; Lakatos, 1978 among others), refutations of scientific theories seem to be as impossible as their definitive confirmation, and furthermore it is not what scientists usually do. Popper himself moved away from this fallibilist position by introducing the concept of corroboration (Lakatos, 1978). Moreover, Klayman and Ha (1987) showed that the strategy called *positive testing* of hypotheses, which can lead to erroneous confirmation in some cases (such as the selection task) is in fact a generally valid strategy, even more than the *negative testing* strategy, which produces the correct response in the selection task.

The second theoretical account of the selection of the *p* & *q* cards is the *matching bias* hypothesis (Evans, 1972a, 1998). According to this concept, the error is not the consequence of a reasoning process, but that of a heuristic one, which consists in the mere selection of the cards mentioned in the rule. In order to test this claim, Evans (1972a) introduced negations in the antecedent and the consequent of the rule, which resulted in the construction of four different selection tasks with the rules:

1. *If p then q*
2. *If p then not-q*
3. *If not-p then q*
4. *If not-p then not-q*

According to the *matching bias* hypothesis most of the participants, in any of these tasks, should select the *p* & *q* cards, irrespective of the presence of the negations.[1]

The third main explanation of the incorrect response in the selection task has been put forward by Johnson-Laird and Wason (1970). In their flow-diagram model, the participants without any insight into the problem would primarily focus on the cards mentioned in the rule. If they consider that the converse of the rule also holds (e.g., *If there is an even number on one side, there is a vowel on the other side*), then they will select the cards *p* & *q*, but will select the card *p* alone if this is not the case. This idea has been refined in the Mental Model theory, one of the most influential theories in the field of reasoning (Johnson-Laird & Byrne, 2002): in the selection task, the participants would only consider the cards that are explicitly represented in their mental model of the conditional rule, and then select those which would matter relative to the truth of the rule. The conditional rule « *If there is an A, then there is a 2* » would flesh out the model:

[A]   2
. . .

(where the square brackets mean that the item is exhaustively represented in the mental model). The participants would take into consideration the two cards "A" and "2", but would then select only the "A" card, which is explicitly represented.

In case of a biconditional reading of the rule, the model would become:

[A]   [2]
. . .

[1]Evans (1984) later revised his account of the selection task based on his Heuristic-Analytic theory, but it remained unchanged for the rule (1) on which we will focus here.

which would lead to the selection of the cards "A" and "2".

The few participants who reach the correct response *p* & not-*q* would elicit the exhaustive model of the conditional:

[A] [2]
[¬A] [2]
[¬A] [¬2]

and then infer the counter-example of the rule (A ¬2).

A psycholinguistic analysis of the selection task in terms of pragmatic inferences has been offered by Sperber, Cara, and Girotto (1995), based on the more general Relevance theory about the pragmatics of language (Sperber & Wilson, 1986). These authors described several processes that might occur in the resolution of the task. One of them is the biconditional interpretation of the rule (inference of the converse) often invited by the context, which along with the application of the Modus Ponens produces the response *p* & *q*. An alternative account of this last selection—favoured by the authors—comes from the fact that in predicate logic, the conditional rule can be stated as $\forall x\ (Px \rightarrow Qx)$, where $\forall$ means "For every" and "$\rightarrow$" the conditional connective. In accordance to the well-known tendency to infer "Some" from "All" (e.g., *All ravens are black* would imply that *There is at least one black raven*), the participants would turn over the *p* & *q* cards in order to show that there is at least one card of the sort described in the rule.

These various explanations of the reasoning processes that could occur in the selection task have rarely been directly challenged within the paradigm of the selection task. However, Hardman (1998) compared the relevance-based theories of Evans (1984) and Sperber et al. (1995) by removing the matching cards. The results lent more support to the latter theory. Here, we focused on the idea that the frequent *p* & *q* pattern of selection is associated with a biconditional interpretation of the IF ... THEN ... rule, as proposed by the Mental Model theory, and evoked as a possible explanation in the Relevance theory.

EXPERIMENT 1

A number of early studies reported some biconditional interpretation in the selection task (Brown, Keats, Keats, & Seggie, 1980; Johnson-Laird & Wason, 1970; Wason, 1968; Wason & Golding, 1974). However, Cheng, Holyoak, Nisbett, and Oliver (1986) proposed explicit biconditionals to their participants such as *If a card has a circle on one side, then it has the word "RED" on the other, and conversely, if it has the word "RED" on one side, then it has a circle on the other*. Whereas the correct response would be to select all four cards, the main error was the omission of *not-p* and *not-q* cards (thus the main response was the selection of the *p* & *q* cards). Platt and Griggs (1993a) brought 69–81% of their participants to select the correct cards, but this result was obtained with the help of notable modifications of the task: the rule was made explicit by adding information (e.g., "A card with a vowel on it can only have an even number, but a card with a consonant on it can have either an even or an odd number."); participants were asked to write some justification of their selection; and the instructions invited them to check whether the rule was being violated, instead of being true or false (violation instructions). Platt and Griggs' modifications of the selection task had, according to them, the effect of blocking the conversion of the rule (in particular the clarification, whereas the response justification and violation instructions could have enhanced analytic reasoning). They noted that their results could be integrated into the framework of the Mental Model theory.

Kirby (1994) also suggested that the correct response *p* & *not-q* was the result of the suppression of the biconditional interpretation of the rule. Green (1997) used the construction task (Evans, 1972b) which consists of asking the participants to complete the cards following the rule, for example, *If a card has the letter A on the top, then it has the number 8 on the bottom*. A vast majority of participants constructed the cards with an "exclusive representation" of the rule—here named biconditional interpretation—by matching the letter A with the number 8 and the letter K with the number 5. The same paradigm was used by Green, Over, and Pyne (1997) with thematic but non-deontic (descriptive) rules, for example, *If this is a day of the week, the manager is in London*. They found that the participants selected more often the *p* & *not-q* cards when they constructed the rule as an implication (and the *p* & *q* cards in the case of an interpretation of the rule as an implication, but this appears to be only a slight, non-significant tendency).

Gebauer and Laming (1997) and Osman and Laming (2001) explicitly advocated that the common pattern in the selection task was due to two "misunderstandings" of the rule. The first one consists in the substitution of a "top/underneath" reading of the rule (*If there is a vowel on top, there is an even number underneath*), and the second one is the biconditional reading of the rule. In their conception, a "top/underneath" reading would lead to the selection of *p* and a biconditional reading to the selection of the four cards, whereas the two misunderstandings would elicit the common *p* & *q* response (*p* & *not-q* being selected when the rule is correctly understood). They tested their hypotheses using a method that could be described as a "partial evaluation task": in six subsequent tasks, participants really had to turn over the cards, and were requested to tell what consequence arose on the truth or falsity of the rule (in the real evaluation task, this has to be done with all the possible occurrences, not only the selected cards). Finally, Feeney and Handley (2000) added a second rule of the sort *if r then q* to the conditional rule *if p then q*, which blocked a biconditional reading of the rule. They observed that this manipulation diminished the frequency of the *p* & *q* cards, but without a concomitant increase of the response *not-q*.

The Wason selection task led to the discovery of a *deontic effect*, the notable facilitation of the problem when presented with a deontic rule such as *If a person is drinking beer, then the person must be over 19 years of age* (e.g., Griggs & Cox, 1982). For deontic versions of the selection task, numerous authors evoked that some bilateral rules (those that can be checked from different perspectives; for example in a supermarket, *If you spend more than £100, then you may take a free gift*, from the point of view of the manager or that of the customers; Manktelow & Over, 1991) are biconditionals when a neutral perspective is taken. Kirby (1994, p. 24) wrote about such rules:

> Cosmides (1989), Gigerenzer and Hug (1992), and Manktelow and Over (1991) have all found problems in which subjects primarily select the not-P and Q cards. However, all of these problems involve deontic conditionals that would normally be interpreted as biconditionals, thus making all four cards logically correct selections.

In an unfamiliar, implausible deontic version of the problem (*If one wants to drive a fluorescent car, then one must drive at a speed over 100 km/h*), Girotto, Gilly, Blaye, and Light (1989) found that 57% of the errors consisted in the selection of the four cards, which is considered by the authors as a "bidirectional" interpretation of the rule. A relatively high rate of the selection of the four cards was observed when a bilateral rule was checked from a neutral point of view (Politzer & Nguyen-Xuan, 1992), and when cheating could be bidirectional (Liberman & Klar, 1996; Light, Girotto & Legrenzi, 1990). Politzer and Nguyen-Xuan (1992) reported 25% of four cards selection, and 29% of *p* & *q* response.

To our knowledge, only one author explicitly disputed the link between the *p* & *q* response and the biconditional interpretation of the rule: Dominowski (1989, in Dominowski, 1995) reported that only 9% of the participants that judged both the *p* & *not-q* and the *not-p* & *q* cards as falsifying at the evaluation task chose to turn over the *p* & *q* cards at the selection task (41% of the latter considered only *p* & *not-q* as falsifying, and 39% all patterns except *p* & *q* in a conjunctive interpretation). But unfortunately these results have to be taken cautiously, as they have not been published.

On the whole, these findings from syllogistic reasoning and conditional inferences strongly suggest the need to assess the interpretation of the rule in the selection task, as it has been done by Green (1997) with the construction task and by Gebauer and Laming (1997) and Osman and Laming (2001) with some kind of "partial evaluation" task. Stenning and van Lambalgen (2004) argued for the need to assess the interpretation participants make in conditional inference tasks. In the present studies, we used a conditional inference task similar to that of Evans (1972b) presented above: for each card presented separately, it was asked what figures on the other side, with the rule being stated as true.

Our first hypothesis in this study is that a deontic rule such as *If a customer is drinking beer, then he/she must be over 18 years of age* adapted from Griggs and Cox (1982) will be considered as a conditional rule (customers who are drinking beer are over 18, customers who are not drinking beer are over or under 18, customers over 18 are drinking beer or Coca-Cola, but customers under 18 only drink Coca-Cola, supposing that the rule is followed). By contrast, a rule such as *If a customer spends more than 100 CHF, then he/she receives a free gift* adapted from Manktelow and Over (1991) will be interpreted biconditionally (customers that spend more than 100.- receive a free gift, customers that spend less than 100 CHF do not receive a free gift, customers that receive a free gift spend more than 100 CHF, but customers that do not receive a free gift spend less than 100 CHF, supposing that the rule is followed).[2] As has been widely claimed, the content of a conditional rule may alter its logical interpretation. Our second hypothesis states that the dominant pattern of choice in the abstract selection task (the *p* & *q* cards) will be linked with a biconditional reading of the IF ... THEN ... rule, whereas the rare "correct" response *p* & *not-q* will be associated with a conditional interpretation of the rule. This would mean that the purported logical "error" or "bias" in this task is in fact the result of a differential interpretation of the rule, and not of faulty reasoning. This hypothesis would also hold for the deontic or realistic versions of the problem.

Method

Participants. A total of 173 first-year psychology students from Geneva University took part in this experiment as part of course requirements.

[2]We slightly modified the original rule "*If you spend more than £100, then you may take a free gift*" because it could suggest that a customer might decide not to take a free gift (*q*) when spending more than £100 (*p*). This would constitute a minor problem for the selection task (removing the modal verb decreases the rate of correct responses, but does not change the general pattern of results, see for example Platt & Griggs, 1993b), but would be problematic for the deduction task.

Materials and procedure. Each participant was randomly assigned to one of four conditions:

1. Bar: *If a customer is drinking beer, then he/she must be over 18 years of age* ($N = 44$)
2. V-E Cards: *If a card has a vowel on one side, then it has an even number on the other side* (Wason & Johnson-Laird, 1972) ($N = 40$)
3. A-3 Cards: *If a card has the letter A on one side, then it has the number 3 on the other side* (Wason, 1968) ($N = 44$)
4. Supermarket: *If a customer spends more than 100 CHF, then he/she receives a free gift* ($N = 45$)

In the first condition, the instructions were as follows (translated from the French):

> Imagine a bar where alcoholic and non-alcoholic drinks are served to customers of all ages. A rule that holds in the bar is written on a board: *If a customer is drinking beer, then he/she must be over 18 years of age*. Imagine that you are a police officer who has to judge if this rule is followed, by checking one or several of the following persons. Which one(s) do you check?

with four persons labelled as "a customer who drinks beer", "a customer who drinks a Coke", "a 22 year-old customer", "a 16 year-old customer". The fourth condition about the supermarket was quite similar, except that no perspective (such as the policeman in the context of the bar) was mentioned.

The material of the second condition was the classical selection task, worded as follows

> Imagine some special double-sided cards: on one side is a letter (vowel or consonant) and on the other side is a number (odd or even). You are given the following rule: *If a card has a vowel on one side, then it has an even number on the other side*. Imagine that you have to judge if this rule is true, by turning over one or several of the cards presented above. Which one(s) do you turn over?

with four cards labelled "E", "K", "4", "7".

The third condition was quite similar, except that the rule was *If a card has the letter A on one side, then it has the number 3 on the other side*, the four cards indicating "A", "B", "3", "2". This version was added in order to test Legrenzi's (1970) idea that the binary nature of the material would favour the biconditional interpretation of the rule.

Along with the selection task, the deduction task was constructed as follows: "About the same rule, *If a card has a vowel on one side, then it has an even number on the other side*, you are given the following cards". First the card "E" was given, then the question was asked, "What figures on the other side (supposing that the rule is true)? (choose): (a) An even number; (b) An odd number; (c) One or the other".

The same was asked with the card "K", and then for the cards "4" and "7": "(choose): (a) A vowel; (b) A consonant; (c) One or the other".

The pattern of response was thus a succession of four letters among three possibilities (a, b, or c), respective to the four answers that were given by the participants. The rates of conditional (pattern "accb") and biconditional ("abab") interpretations of the rule were computed, the remaining patterns were classified as "other" interpretations.

In the Bar and the Supermarket versions of the problem, the deduction task was similarly constructed, regarding customers respectively in the bar and the supermarket.

For each participant, the same rule was used in the selection task and in the deduction task, in order to avoid possible interferences between the interpretations of the rule. The order of the two problems (selection task and deduction task) was balanced in each experimental condition.

Results

First, the order of the problems had no effect on the responses, neither at the deduction task, $\chi^2(2, N = 173) = 1.51$, *ns*, nor at the selection task, $\chi^2(2, N = 173) = 1.86$, *ns*. The two orders are thus combined in the following analyses.

The deduction task leads to the following results (see Figure 1): The frequency of patterns differs by experimental condition, $\chi^2(6, N = 173) = 102.3$, $p < .0001$, Cramer's $V = .54$. The bar rule is clearly interpreted as a conditional rule (75%), but the supermarket rule as a biconditional rule (91%). These results strongly support the first hypothesis. Abstract rules are more ambiguous, mostly interpreted as biconditionals (V-E: 60% of biconditional pattern versus 13% of conditional pattern and 27% of "other" patterns; A-3: 48% of biconditional pattern versus 16% of conditional pattern and 36% of "other" patterns). Although consistent with the hypothesis inspired by Legrenzi (1970), the difference between the two abstract versions is not significant, $Z = 1.10$, *ns*.[3]

The results of the selection task are summarised in Figure 2 (less-frequent patterns are aggregated). The frequency of patterns differs relative to the context evoked in the problem, $\chi^2(6, N = 173) = 67.08$, $p < .0001$, Cramer's $V = .44$. From the classical point of view (the rule is interpreted as a conditional rule and the correct response is p & *not*-q), the highest rate of

[3]The z-statistic refers to the proportion test (e.g., Kanji, 1993), which allows testing of a one-tailed difference of proportions (unlike the chi-square). The latter will be used when no precise hypotheses are made.

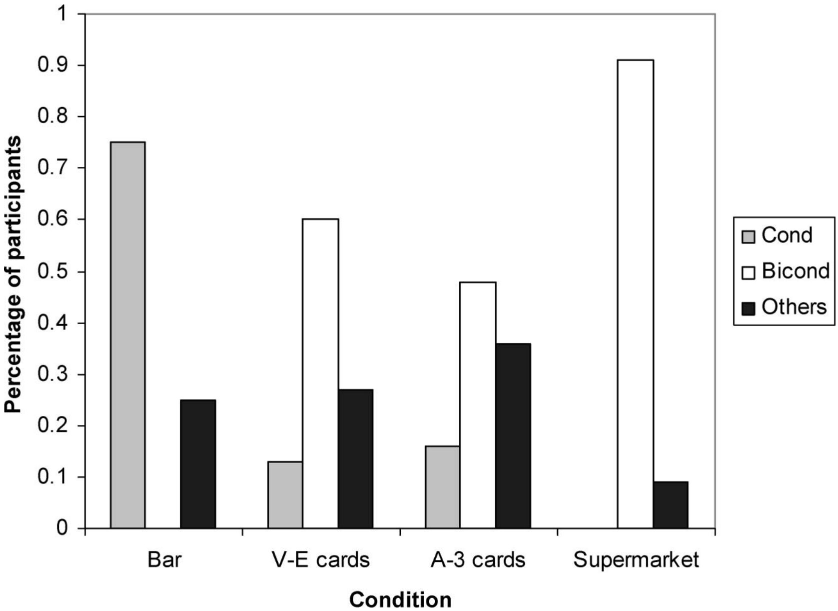

Figure 1. Percentage across condition in the deduction task (cond = conditional interpretation; bicond = biconditional interpretation).

success is observed in the bar condition (41% of correct responses, versus 8% in the V-E condition, 5% in the A-3 condition, and 7% in the context of the supermarket). This rate is nonetheless lower than in Griggs and Cox's (1982) study, but this could be due to the instructions that invite participants to judge if this rule is followed, whereas traditional instructions ask
participants to judge if the rule is violated. For example, in a replication of 1
Yachanin (1986), Dominowski (1990, in Dominowski, 1995) reported a drop in the rate of correct responses from 87% to 46% when such a change in the instructions was made. Correlatively, 45% of the participants in the bar condition only choose the card *p*. In the supermarket problem, the responses are more diverse, with a predominance of *q* (33%) and *p* & *q* (18%) choices. The abstract selection tasks lead to the classical results: a predominance of responses *p* & *q* (V-E: 57%, A-3: 59%) and *p* alone (V-E: 19%, A-3: 20%).

In order to evaluate the second hypothesis, the link between the two tasks was investigated, with all conditions combined (see Table 1). It was statistically significant, $\chi^2(4, N = 173) = 36.72$, $p < .0001$, Cramer's $V = .33$ (*abab*, *accb*, and other patterns across *pq*, *p*¬*q* and other patterns). As expected, the great majority of the *p* & *q* responses (67%) are associated

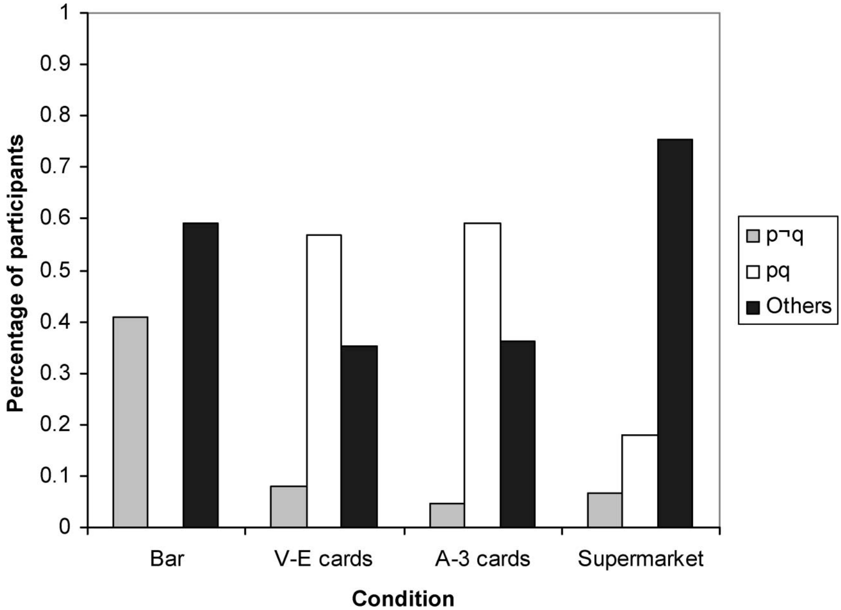

Figure 2. Percentage across condition in the selection task.

TABLE 1
Frequency of responses across the two tasks (Exp. 1)

	Pattern				
Selection	*abab*	*accb*	*acac*	*Others*	*Total*
pq	37	4	12	2	55
p	11	18	1	12	42
p¬*q*	4	18	0	4	26
q	13	0	0	3	16
p¬*pq*¬*q*	8	0	0	2	10
Others	13	5	0	6	24
Total	86	45	13	29	173

abab = biconditional interpretation; accb = conditional interpretation.

with the biconditional pattern, whereas a smaller proportion of them (22%) relates to the pattern *acac* (which can be viewed as something close to the conjunctive interpretation of conditionals, Barrouillet & Lecas, 2002). The selection of the *p* & *not*-*q* cards is mostly linked to the conditional pattern (69% of *p* & *not*-*q* selections), and far less to the biconditional pattern

(15%). The second hypothesis is thus clearly supported. Participants who select the card *p* are divided between those who interpret the rule as a conditional (43%) and those who show a biconditional reading of the rule (26%). The choice of the card *q* (mostly observed in the context of the supermarket) and the selection of all four cards are tightly associated with a biconditional interpretation of the rule (for respectively 81% and 80% of the participants).

In the other direction, participants who produce the biconditional pattern choose to select mainly *p* & *q* (44%) in the selection task, and also less frequently *p* (13%), *q* (15%), and all four cards (10%). The conditional pattern is associated with the selection of *p* & *not*-*q* (for 40% of the participants) or *p* only (for another 40% of the participants).

Our second hypothesis had to be more specifically tested in the abstract versions of the selection task (V-E and A-3 ; see Table 2). In these two conditions, 62% of the participants who select *p* & *q* consider the rule as a biconditional (and 26% as a conjunction *acac*), whereas 60% of the responses *p* & *not*-*q* are related to a conditional interpretation (and 20% to the biconditional interpretation). As previously, the hypothesis is corroborated.

Conversely, 67% of the participants who show a biconditional reading of the rule turn over the cards *p* & *q* (12% the card *p* and 9% all four cards). A total of 33% of the few participants who interpret the rule as a conditional rule (13–16%), nonetheless select the cards *p* & *q* in the Wason task, *versus* 25% who gave the correct answer *p* & *not*-*q*, and 25% the answer *p*.

EXPERIMENT 2

The second experiment aimed, besides replication, at explaining why most participants only turn the *p* & *q* cards when the rule is considered as a biconditional, while all four cards are relevant in this case. One possible

TABLE 2
Frequency of responses across the two tasks in the abstract versions (Exp. 1)

	Pattern				
Selection	*abab*	*accb*	*acac*	*Others*	*Total*
pq	29	4	12	2	47
p	5	3	1	7	16
$p\neg q$	1	3	0	1	5
$p\neg pq\neg q$	4	0	0	1	5
Others	4	2	0	2	8
Total	43	12	13	13	81

abab = biconditional interpretation; accb = conditional interpretation.

reason could be that from a pragmatic point of view, the framing of the problem that asks someone to turn some cards within an array certainly implies that not all cards should be turned over, unless the task is a real trap. This phenomenon would be comparable to the fact that in some Piagetian tasks asking the same question twice induces children to think that something has changed in the problem, and thus that the answer should also vary (Donaldson, 1978). Imagine somebody showing two objects and asking you: "Which one do you want to take?" It is pragmatically unlikely that you will take them both (unless you are a skinflint or have a good sense of humour). Such a pragmatic inference seems close to one of Grice's (1975) central claims that utterances automatically create expectations which guide the hearer towards the speaker's meaning.

The hypothesis is that in case of biconditional interpretation of the rule in the deduction task, the participants will more often select the four cards in Wason's selection task when they are presented successively rather than together as in the standard problem. However, we will not predict a complete disappearance of the common response *p* & *q*, as a successive presentation will not, in our opinion, obviate the pragmatic inference which suggests that not all cards must be turned over—but it should at least weaken it.

Method

Participants. A total of 76 first-year psychology students from Fribourg University took part in the experiment as part of course credits.

Materials and procedure. Three of the four experimental conditions of the first study were run in three independent groups:

1. Bar: *If a customer is drinking beer, then he/she must be over 18 years of age* ($N = 26$)
2. Cards: *If a card has a vowel on one side, then it has an even number on the other side* ($N = 25$)
3. Supermarket: *If a customer spends more than 100 CHF, then he/she receives a free gift* ($N = 25$)

The measures were identical to those in the first experiment, except that in the selection task the four cards were presented successively and accompanied by the question: "Do you turn it over?" (with two possible answers, "Yes" or "No"). This procedure had already been used by Thompson (2000) for other purposes, in a large variety of descriptive rules. As the order of presentation of the two problems had no effect in Experiment 1, it was kept constant in this study (first the selection task followed by the deduction task).

Results

The following results were obtained in the deduction task (see Figure 3): The frequency of the patterns observed differs according to the version of the problem, $\chi^2(6, N = 76) = 52.66$, $p < .0001$, Cramer's $V = .59$. The results of Experiment 1 are replicated, as the rule in the context of the bar is mainly considered as conditional (85%), whereas the rule in the context of the supermarket is viewed as a biconditional (96%). As previously, the abstract rule lies in between, with 44% of biconditional *versus* 32% of conditional interpretations. The unexpected increase of the conditional interpretation relative to Experiment 1 is nevertheless nonsignificant, $\chi^2(1) = 2.36$, *ns*.

The results of the modified selection task are summarised in Figure 4: The version of the problem significantly alters the patterns of card selection, $\chi^2(6, N = 76) = 29.54$, $p < .0001$, Cramer's $V = .44$. The bar problem elicits a larger rate of correct conditional response (65% of *p* & *not-q* selections) than the other versions (cards: 22%, supermarket: 1%). The selection of the four cards becomes the dominant pattern in the latter problem (29%), whereas the choice of *p* & *q* remains important (21%). The abstract task still generates the responses *p* & *q* (30%), *p* & *not-q* (22%), but also the selection of the four cards (13%).

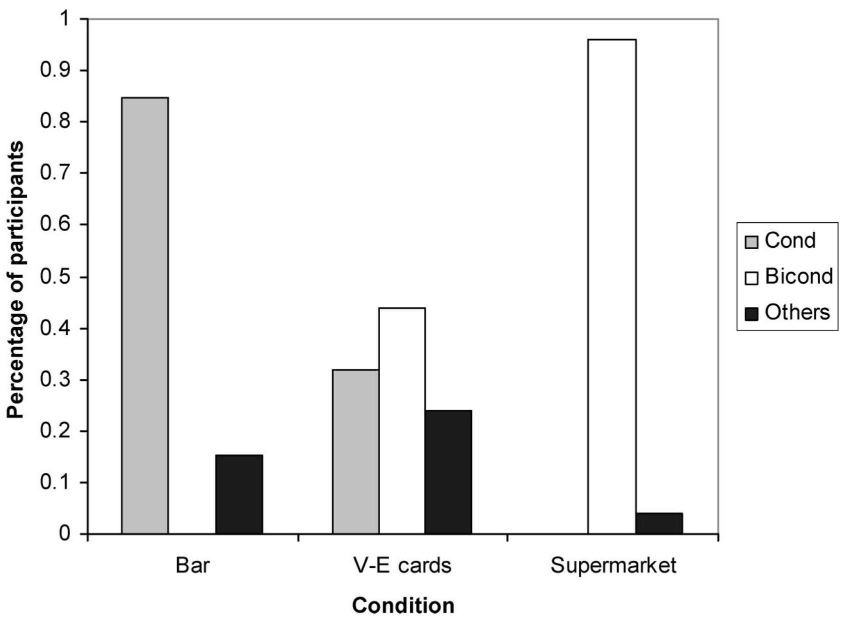

Figure 3. Percentage across condition in the deduction task (cond = conditional interpretation; bicond = biconditional interpretation).

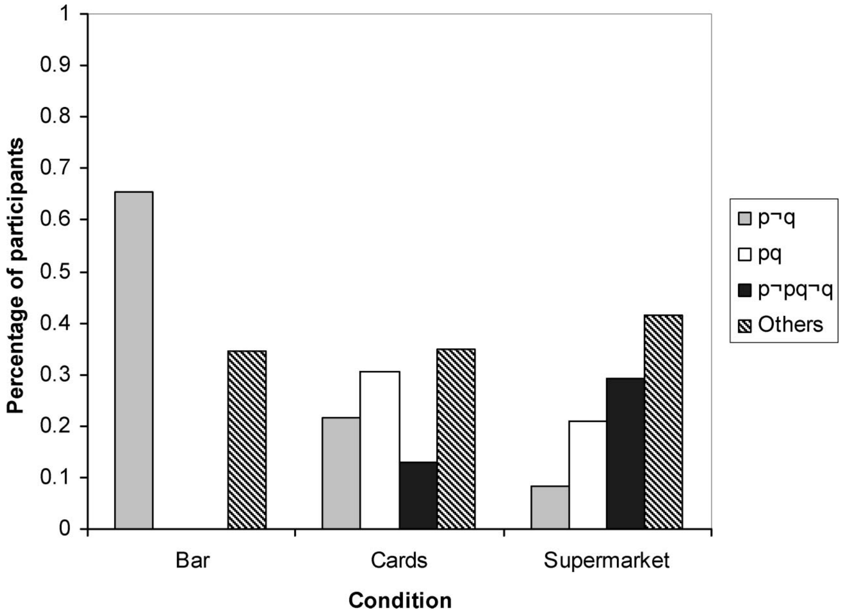

Figure 4. Percentage across condition in the modified selection task.

Compared to Experiment 1, the modified selection task led to an increase in the *p* & *not-q* response rate (65% versus 41%) in the bar scenario, which is closer to the usual rate (73% in Griggs & Cox, 1982), and a correlative decrease in the *p* response rate (from 45% to 8%). As expected, the rate of selection of all four cards increased, especially in the problem of the supermarket (29% *versus* 11% in Experiment 1) but not markedly in the abstract task (13% *versus* 8%). Nevertheless, it must be kept in mind that in the latter condition, the biconditional interpretation was somewhat less frequent than in Experiment 1.

With all conditions combined, the link between the two tasks was investigated (see Table 3). It was statistically significant, $\chi^2(6, N = 76) = 34.23$, $p < .0001$, Cramer's $V = .56$ (*abab*, *accb*, and other patterns across *pq*, *p¬q*, *p¬pq¬q* and other patterns). As in Experiment 1, the *p* & *q* response relates mainly to the biconditional interpretation for 67% of the participants, and a little (17%) to the conjunctive interpretation (pattern *acac*). The correct conditional response *p* & *not-q* is as strongly associated with the conditional reading of the rule for 79% of the participants. All of the 10 persons who chose to turn over the four cards consider the rule as biconditional in the deduction task. Conversely, most participants who interpret the rule biconditionally turn

TABLE 3
Frequency of responses across the two tasks (Exp. 2)

	Pattern			
Selection	*abab*	*accb*	*Others*	*Total*
$p\neg q$	2	19	3	24
pq	8	1	3	12
$p\neg pq\neg q$	10	0	0	10
Others	15	10	5	30
Total	35	30	11	76

abab = biconditional interpretation; accb = conditional interpretation.

over the four cards (29%) and slightly fewer *p* & *q* (23%), whereas the majority of those who treat the rule as a conditional turn over the cards *p* & *not-q* (63%).

In order to test the hypothesis that predicts an increase of the selection of the four cards due to their sequential presentation, we had to take into consideration the participants who show a biconditional interpretation of the rule (and not the simple frequency of card selection). The proportion of those participants rises from 10% in the first experiment to 29% here, $Z = 2.60$, $p < .004$, for both the supermarket problem (from 10% to 29%) and the abstract version (from 9% to 27%), in accordance with our hypothesis.

The rate of the *p* & *q* response jointly decreases from 44% to 23%, $\chi^2(1, N = 121) = 4.72$, $p < .03$, which makes the selection of the four cards the dominant pattern of selection in cases of a biconditional interpretation of the rule. Each card being taken separately, the participants in the first study far more often select the *p* (96%) and *q* (87%) cards than *not-p* (15%) and *not-q* (19%) cards, but this is not the case in Experiment 2 (*p*: 75%, *q*: 69%, *not-p*: 56%, *not-q*: 50%).

GENERAL DISCUSSION

Our first results have illustrated a common observation in the psychology of reasoning, which asserts that the "logical" structure of IF ... THEN ... expressions in terms of conditional versus biconditional depends on the meaning of the propositions and/or the context in which the rule takes place. We saw how rules like *If a customer is drinking beer, then he/she must be over 18 years of age*, in the context of a bar, and *If a customer spends more than 100 CHF, then he receives a free gift*, in the context of a supermarket, were quite unambiguously interpreted respectively as a conditional and as a biconditional by the participants in Experiments 1 and 2. On the other hand, abstract rules about cards such as *If a card has a vowel on one side, then it has* *an even number on the other side* are more ambiguous, although they were, as previously suggested (Evans et al., 1993), mainly interpreted as biconditional rules.

In the introduction of this paper, we summarised the discussion that occurred among psycholinguists and cognitive psychologists about the correctness of the conversion (or invited inference) of the rule *if p then q* to *if q then p*. Despite the continuing and inappropriate use of pejorative terms like "bias" or "fallacy", the conclusion is clearly that, from a pragmatic point of view, the conversion is perfectly reasonable. We could go even further by claiming that "biconditional interpretation" should be preferred to concepts like the "conversion" of the rule, "invited inferences", or "misinterpretation of premises", all of which imply that the conditional interpretation (as defined in formal logic) of IF... THEN... expressions is the correct one. On the contrary, it seems that formal logic, whose aim was to deal with universal (thus abstract, devoid of any content) statements, had to solve the apparent ambiguity of language by other means, which is typically clarified in a pragmatic way by the content and context of speech (as with the exclusive/inclusive disjunction for example). This was done by creating an artificial connective, the biconditional *if and only if p then q*, which is indeed absent in natural language. It seems then quite ironic that pragmatic inferences could be judged as erroneous from the point of view of formal logic (see Cohen, 1981, and Moscovici, 1976, for similar views).

The present results and other studies (Gebauer & Laming, 1997; Green, 1997; Osman & Laming, 2001) have shown that the modal response in the selection task (*p* & *q* cards) is mostly linked with a biconditional interpretation of the rule, as had been hypothesised by Johnson-Laird and Wason (1970) and developed later in Mental Model theory (Johnson-Laird & Byrne, 2002). Thus, in our view, it should not be considered as some kind of bias or error committed by the majority of the participants (whether we call it confirmation or matching bias), as is still argued nowadays in most textbooks in cognitive psychology. Rather, it should be considered as a restriction of one's choice (the correct answer in the case of a biconditional interpretation of the rule is to select all four cards). This limitation hardly constitutes an error, because of the pragmatic demands of the task. Our point of view is that an utterance such as "Which one(s) do you turn over?" concerning four cards pragmatically implies that all cards should not be turned over, unless one suspects a trick, which is rather improbable (though certainly not impossible!) in a psychological experiment. The results of Experiment 2 give partial support to this interpretation. However, other tests of this hypothesis should be carried out: instructions could be modified as "Select every card that could violate the rule", following Griggs and Jackson (1990). Other explanations of the limitation of card selection can be derived from existing theories, given that they incorporate the prior interpretational stage in terms of conditional/biconditional: matching strategy (but not bias because of the biconditional interpretation), probabilistic factors (Green et al., 1997; Kirby, 1994; Yama, 2001), information gain (Oaksford & Chater, 1994, 2003), mental models, relevance processes, etc. What is clearer is that the observed link between interpretation and selection is not compatible with purely non-logical accounts of the selection task, as previously advocated by numerous authors (Feeney & Handley, 2000; Johnson-Laird & Byrne, 2002; Platt & Griggs, 1993a; Sperber et al., 1993). (1)

Although the conditional/biconditional interpretation seems to bear some importance in the selection task, a substantial part of the participants gave relatively idiosyncratic or rare patterns of selection and inference. This is an illustration of the seemingly numerous interpretational problems that arise in the abstract selection task, as Stenning and van Lambalgen (2004) thoroughly investigated.

Concerning deontic versions of the selection task, we observed that the absence of a perspective taken on bilateral rules led to a biconditional interpretation of such rules (as in Girotto et al., 1989; Liberman & Klar, 1996; Light et al., 1990; Politzer & Nguyen-Xuan, 1992). The present results thus seem compatible with the idea that the introduction of a point of view on bilateral rules leads one to focus on one direction: *If a customer spends more than 100 CHF, then he/she must receive a free gift* from the perspective of the customers (*if p then q*, checking for the *p* & *not-q* cases), and *If a customer receives a free gift, then he/she must have spent more than 100 CHF* for the store manager (*if q then p*, checking for *q* & *not-p* cases). This interpretation is in line with several previous analyses (Holyoak & Cheng, 1995; Kirby, 1994; Liberman & Klar, 1996; Sperber et al., 1995; Thompson, 2000; Wagner-Egger, 2001).

In conclusion, all the intensive work on Wason's selection task illustrates well the need to consider analytic and heuristic factors, as it has been emphasised by Evans and Over (2004), but also pragmatic factors in the entire reasoning process. Whereas the analytic side of reasoning has been extensively studied, some authors have stressed the need to further explore the interpretative part of reasoning (Noveck & Sperber, 2004; Politzer, 1997; Sperber et al., 1995; Thompson, 2000). One interesting conclusion of the two present studies is that the same interpretational step in terms of conditional/biconditional seems to affect both abstract and deontic versions of the selection task.

Manuscript received 8 November 2005
Revised manuscript received 10 April 2007
First published online day/month/year

REFERENCES

Barrouillet, P., & Lecas, J-F. (2002). Content and context effects in children's and adults' conditional reasoning. *Quarterly Journal of Experimental Psychology*, *55A*(3), 839 – 854.

Braine, M. D. S., & O'Brien, D. P. (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Brown, C. A., Keats, J. A., Keats, D. M., & Seggie, I. (1980). Reasoning about implication: A comparison of Malaysian and Australian subjects. *Journal of Cross-Cultural Psychology*, *11*, 395 – 410.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, *58*, 220 – 226.

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

Cohen, L. J. (1981). Can human rationality be experimentally demonstrated? *The Behavioral and Brain Sciences*, *4*, 317 – 370.

Dominowski, R. L. (1995). Content effects in the selection task. In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Donaldson, M. (1978). *Children's minds*. Glasgow, UK: Fontana-Collins.

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

Evans, J. St. B. T. (1972b). Reasoning with negatives. *British Journal of Psychology*, *63*, 213 – 219.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, *75*, 451 – 468.

Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: Do we understand it after 25 years? *Thinking & Reasoning*, *4*, 45 – 82.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, UK: Oxford University Press.

Feeney, A., & Handley, S. J. (2000). The suppression of q card selections: Evidence for deductive inference in Wason's selection task. *Quarterly Journal of Experimental Psychology*, *53A*, 1224 – 1242.

Fillenbaum, S. (1975). IF: Some uses. *Psychological Research*, *37*, 245 – 260.

Gebauer, G., & Laming, D. (1997). Rational choices in Wason's selection task. *Psychological Research*, *60*, 284 – 293.

Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry*, *11*, 561 – 565.

Girotto, V. (1990). Biases in children's conditional reasoning. In J-P. Caverni, J.-M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases*. Amsterdam: Elsevier.

Girotto, V., Gilly, M., Blaye, A., & Light, P. H. (1989). Children's performance in the selection task: Plausibility and familiarity. *British Journal of Psychology*, *80*, 79 – 95.

Green, D. W. (1997). Hypothetical thinking in the selection task: Amplifying a model-based approach. Commentary on "Rationality in reasoning: The problem of deductive competence" by J. St. B. T. Evans & D. E. Over. *Current Psychology of Cognition*, *16*, 93 – 101.

Green, D. W., Over, D. E., & Pyne, R. A. (1997). Probability and choice in the selection task. *Thinking & Reasoning*, *3*, 209 – 235.

Grice, P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Studies in syntax, Vol. 3: Speech acts*. New York: Academic Press.

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

Hardman, D. (1998). Does reasoning occur on the selection task? A comparison of relevance-based theories. *Thinking & Reasoning*, *4*, 353–376.

Holyoak, K. J., & Cheng, P. W. (1995). Pragmatic reasoning with a point of view. *Thinking & Reasoning*, *1*, 289–313.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, *1*, 134–148.

Kanji, G. K. (1993). *100 statistical tests*. London: Sage.

Kirby, K. N. (1994). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition*, *51*, 1–28.

Klayman, J., & Ha, Y-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge, UK: Cambridge University Press.

Legrenzi, P. (1970). Relation between language and reasoning about deductive rules. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistic*. Amsterdam: North-Holland.

Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: Social exchange cheating detection or task understanding? *Cognition*, *58*, 127–156.

Light, P. H., Girotto, V., & Legrenzi, P. (1990). Children's reasoning on conditional promises and permissions. *Cognitive Development*, *5*, 369–383.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, *39*, 85–105.

Matalon, B. (1962). Étude génétique de l'implication. In J. Piaget (Ed.), *Études d'épistémologie génétique*. Paris: Presses Universitaires de France.

Moscovici, S. (1976). *La psychanalyse, son image et son public*. Paris: Presses Universitaires de France.

Noveck, I., & Sperber, D. (Eds.). (2004). *Experimental pragmatics*. London: Palgrave.

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. (2003). Conditional probability and the cognitive science of conditional reasoning. *Mind & Language*, *18*(4), 359–379.

Osman, M., & Laming, D. (2000). Misinterpretation of conditional statements in Wason's selection task. *Psychological Research*, *65*, 128–144.

Platt, R. D., & Griggs, R. A. (1993a). Facilitation in the abstract selection task: The effect of attentional and instructional factors. *Quarterly Journal of Experimental Psychology*, *46A*, 591–613.

Platt, R. D., & Griggs, R. A. (1993b). Darwinian algorithms and the Wason selection task: A factorial analysis of social contract selection task problems. *Cognition*, *48*, 163–192.

Politzer, G. (1997). Rationality and pragmatics. Commentary on J. St. B. T. Evans & D. E. Over "Rationality in reasoning: The problem of deductive competence". *Current Psychology of Cognition*, *16*, 190–195.

Politzer, G., & Nguyen-Xuan, A. (1992). Reasoning about conditional promises and warnings: Darwinian algorithms, mental models, relevance judgments or pragmatic schemas? *Quarterly Journal of Experimental Psychology*, *44*, 401–412.

Popper, K. (1959). *The logic of scientific discovery*. London: Routledge & Kegan Paul.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *57*, 31–95.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Stenning, K., & van Lambalgen, M. (2004). A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, *28*, 481 – 529.

Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, *76*, 209 – 268.

Wagner-Egger, P. (2001). Costs and benefits in Wason's selection task: The Social Conditional Model. *Swiss Journal of Psychology*, *60*, 117 – 135.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.) *New horizons in psychology I*. Harmondsworth, UK: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273 – 281.

Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*, 141 – 154.

Wason, P. C., & Golding, E. (1974). The language of inconsistency. *British Journal of Psychology*, *65*, 537 – 546.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Harmondsworth, UK: Penguin.

Wilkins, M. C. (1928). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, *102*.

Yama, H. (2001). Matching versus optimal data selection in the Wason selection task. *Thinking & Reasoning*, *7*(3), 295 – 311.