Dual processes in reasoning?*

P. C. WASON

University College London

J. ST. B. T. EVANS

Plymouth Polytechnic

Abstract

Previous results have shown that the introduction of negatives into the sentences used in a deductive problem affected behavior in a systematic way which was independant of the logical structure of the problem. In the present investigation, the subjects were asked to justify their responses when reasoning about such sentences. In accordance with previous results, the responses were dominated by the terms in the sentences regardless of whether they were negated. However, the justifications did vary when negatives were introduced in accordance with the logical consequences of the responses. The interpretation of these justifications as causes of behavior seemed implausible. It was suggested that they were rationalizations, or that there was at least some form of dual processing between behavior and conscious thought.

The aim of this investigation is to elucidate the character of the reasons given for attempted solutions to the *selection task*, or *four-card problem* (Wason, 1966, 1968a). The problem is now fairly well known. In essence it consists in establishing the truth value of a conditional sentence, e.g., 'If a card has a vowel on one side, then it has an even number on the other side', by selecting for inspection the necessary and sufficient cards from a set consisting of a vowel, a consonant, an even number and an odd number (under the restriction that each has a letter on one side and a number on the other side). The solution is to select the vowel and the odd number because only these two

* We thank I. Keddie for conducting the experiment, and Dr. Neal Smalley for invaluable criticism. First author's address:

Psycholinguistics Research Unit, Wolfson House, 4, Stephenson Way, London N.W. 1 2HE.

values on the same card could falsify the test sentence; the even number could merely verify it vacuously. But previous studies (with abstract material) have demonstrated that the most frequent erroneous solutions consist in selecting the values mentioned in the test sentence. The problem is recalcitrant to correction (Wason 1969a), and its difficulty is not restricted to test sentences with the surface structure of a conditional; in a recent version (Wason and Golding, 1974) the sentence is an assertion e.g., 'a letter is above each number'.

The reader might well expect some justification for yet another paper on the problem. The issues suggested by previous research still remain puzzling and have only recently been investigated systematically. For instance, Bracewell (1974) has argued that the subjects' remarks, when confronted by subsequent contradictions to their initial choice, (e.g., 'That is doing it the other way round') are 'correct' inferences based on inadequate premises: The subject is supposed to lack access to all the implicit features of the problem. Our view is that such utterances are rationalizations determined by erroneous selections. The issue is of some generality, and the present study shifts the argument to the character of elicited reasons which accompany the choice of cards. The problem has also been turned to other purposes, e.g., a critique of formal operations (Wason, 1975), the logic of natural language (Van Duyne, 1974) and interhemispheric differences in reasoning (Golding, Reich and Wason, 1974). It has, in fact, become a tool for investigating a variety of theoretical issues rather than an object of study in its own right.

An information-processing model (Johnson-Laird and Wason, 1970) has been postulated as an explanation of performance. A subject's degree of insight into the problem is defined as a function of his appreciation that falsification of the test sentence is more relevant to the solution than its verification. Goodwin and Wason (1972) found a correlation between the postulated degrees of insight (defined by choice of cards) and reasons given for the choice.

Evans and Lynch (1973), however, adduced evidence that error is due not to a tendency to verify but to a tendency to 'match' the mentioned cards with the actual cards. When the consequent of the conditional was negated, they observed a highly significant tendency for a subject to be correct for apparently spurious reasons. For example, given the sentence, 'if there is a W on one side of a card, then there is not a 9 on the other side', the W and the 9 tended to be (correctly) selected. A number which is not 9 available, as (say) 8, is (correctly) rejected. This result might seem to indicate that negating the consequent allows the subjects to be 'right' because they simply ignore the 'not'. The phenomenon, however, is more general. When the antecedent was negated, they observed a highly significant tendency for the value mentioned to be (wrongly) selected. Given the sentence, 'If there is not a T on one side of a card, then there is a 5 on the other side', the T and the 5 tended to be selected although the correct solution would be to select the letter other than T and the number other than 5.

Evans (1972b) argues that matching bias is a hypothesis based on a statistical regularity and that it interacts with the logical processes demanded by the problem. Its lack of theoretical status has been challenged by Van Duyne (1973). But whatever its interpretation, the results obtained by Evans and Lynch (1973) cast some doubt on the Johnson-Laird and Wason (1970) model. Verification and matching coincide (under affirmative sentences), but matching is evidently more fundamental because it appears indifferent to the presence of a negative. The thought processes elicited by this problem seem to be more primitive than had been supposed. Evans and Lynch did not, however, ask their subjects to give reasons for their solutions. The present study aims to see how solutions are related to reasons when the consequent of the conditional is negated. On this particular rule matching bias should induce the logically correct solution, but obviously for the wrong reasons. It is a matter of considerable interest to see whether the subjects' justifications still reflect the 'complete insight' observed by Goodwin and Wason (1972) to accompany the correct solutions.

Experiment

Performance under a test sentence with an unnegated consequent was compared with that under a negated consequent. The sentences had the following general form:

- (1) 'If there is a [specified letter] on one side of a card, then there will be a [specified number] on the other side.'
- (2) 'If there is a [specified letter] on one side of a card, then there will not be a [specified number] on the other side.'

The four cards which accompany each test sentence have the following general form: The specified letter (L+); the specified number (N+); an unspecified letter (L-); and an unspecified number (N-).

Under (1) the solution is (L+) and (N-); under (2) it is (L+) and (N+). It was predicted that values of (L+) and (N+) would be selected under both (1) and (2), the solutions being nominally wrong under (1) and nominally right under (2).

Reasons for either selecting or rejecting each card were written down by the subjects, but no predictions were made about their content.

A related group design was used. Twenty-four unpaid volunteer students of

the City of London Polytechnic, who had no previous experience with the problem, were assigned alternately to one of two groups and tested individually. One group first performed the 'affirmative task' and then the 'negative task'. The other group performed the tasks in the opposite order.

The subjects were first acquainted with the fact that all the cards had a letter on one side and a number on the other side by inspection of an 'example set' of eight cards which were not actually used in the experiment. Each test sentence was presented with four cards displaying (respectively) the two specified and two unspecified symbols. For each subject all the symbols were different on the two tests, and the order of presenting the cards from left to right was random. After the subjects had examined the 'example sets' of cards, the following instructions were read by them, and they were then asked whether they had any questions. Finally, the salient points were repeated by the experimenter to clarify the procedure.

'You will be given two sheets of paper one at a time. At the top of the sheet is stated a simple rule connecting the combinations of letters and numbers which are written on the cards. The rule only applies to the cards in front of you and it may be true or false. You must decide which card (or cards) need definitely to be turned over in order to establish whether the rule is true or false.

Below the cards, on the left-hand side of the sheet, the identification numbers of the cards are arranged in columns. When you have decided which cards must be turned over, place a "Yes" or "No" in the column next to the appropriate identification number:

"Yes" – for turning the card over

"No" - for leaving the card alone

I also want you to write down your reasons for choosing to examine or to ignore *each* card in the third column.

The problem is not so easy as it looks, so please think carefully before giving your answer.'

No time limits were imposed, and the subjects were not allowed to turn over any of the cards.

Quantitative results

The frequency of different selections is presented in Table 1. It will be noted that there is such strong support for the prediction that statistical evaluation is unnecessary. Fifteen out of the 24 subjects were correct in the negative task compared with none in the affirmative task. Furthermore, exactly half the subjects made the matching response (L+) and (N+) in the affirmative task, out of a possible 15 responses. It will also be noted that there is no apparent order effect. However, there were 8 out of 24 atypical responses ('others') defined as selections which either omit (L+) or include (L-), in the affirmative task, compared with three in the negative task (of these 11 cases, four consisted in the selection of all four cards). Previous research (Johnson-Laird and Wason, 1970) suggests that atypical responses are probably indicative of either misunderstanding or guessing. No reason is apparent for the relatively greater number of such responses in the affirmative task.

	Affirm	native	Negative			
Order of permormance/	1st	2nd	Total	1st	2nd	Total
values selcted						
(L+)(N+)	6	6	12	9	6	15*
(L+)	2	2	4	2	4	6
(L+)(N-)	0	0	0*	0	0	0
others	4	4	8	1	2	3
N	12	12	24	12	12	24

Table 1. Frequency of responses in affirmative and negative tasks

Qualitative results

A small minority of protocols were so idiosyncratic that no consistent thought process could be inferred from them. An attempt was first made to classify the remainder with respect to selected cards into the following categories: (a) Verification, (b) falsification and (c) matching. These exemplars illustrate the criteria for assigning a protocol to each category.

- (a) Verification. 'Should be J on the other side if rule holds.' 'If there is not a 3 on the other side the rule is proved.' 'To verify the rule as to there being a B on the opposite side.' 'If there is an A on the other side the statement is true.'
- (b) Falsification. 'If there is a 6 on the other side, then the statement is false.' 'To see if there is a Y on the other side to disprove the statement.' 'If rule is true, then there will not be a 2. If rule is false there will be.' 'If this card is overturned and there is an F the statement is false.'
- (c) Matching. 'Uppermost symbols may correspond with those on the back.' 'It would be possible for there to be C on the other side.' 'If it has a 4

^{* =} Correct response.

overleaf then Q and 4 are associated.' The rule only says that 7 is related to J.,

Some difficulties were experienced in attempting to distinguish between verification and matching. For example, 'To find out if there is a 4 on the other side' could be interpreted as a weak statement of verification. Similarly, 'To check the truth or falsity of the statement' seems like a simple compliance with the instruction. The source of the difficulty is that terms indicative of verification are linguistically unmarked. This difficulty does not arise with the falsification category because its terms are linguistically marked, e.g., 'untrue', 'disproof' or cognate expressions. Hence a more stringent decision was adopted: Each protocol was classified as either 'falsification' or 'not falsification'.

In the affirmative task four out of the 24 protocols were classified as 'falsification'. In the negative task 11 out of 24 were classified as falsification, nine of which were associated with correct solutions. Of these nine, eight occurred in the group who performed the negative task first. This significant order effect (p < 0.05, Fisher exact test) may have been due to the fact that the subjects were not allowed to turn over the cards. This may have induced frustration and hence diminished interest for the second task although there was no evidence for this supposition.

The remaining two subjects, who yielded falsification protocols but incorrect solutions in the negative task, selected in each case just (L+), one in the first performance and the other in the second performance. These cases reveal only an 'indirect falsification effect', i.e., a reason in terms of falsification to justify the selection of the (unnegated) antecedent when the consequent is negated, e.g., 'Easiest thing to do to prove the rule false is to turn T over to find out whether it is in fact also 5.' This is distinguished from a 'direct falsification effect', i.e., a similar type of reason to justify the selection of the (negated) consequent.

Of the four subjects who yielded falsification protocols in the affirmative task, three did so when that task was performed second, and all four yielded fasification protocols in the negative task.

The protocols of three subjects show how the reasons given for the choice of the mentioned cards are quite different in the affirmative and negative tasks when both make exactly the same matching responses. In order to ease the readers' comprehension, the affirmative and negative sentences have been corrected so that they are lexically equivalent. It will be noted that all three cases reveal both a 'direct' and an 'indirect' falsification effect. In addition, the reasons given for not selecting the unmentioned cards reveal a process of 'mismatch'.

Cards and responses		Negative task	Affirmative task		
		If there is a B then there will not be a 3	If there is a B then there will be a 3.		
G 2,	S 3	Reasons	Reasons		
В	yes	'If the rule was false, there would be a 3 on the other side. If true there would not be a 3. B and 3 should be taken as part of the same assumption.'	'The rule only says that B is related to 3. It does not say anything about there being a logical sequence of letters to numbers, so no assumptions about letters and numbers other than B and 3 can be made.'		
3	yes	'If the rule was false, there would be a B on the other side.'	'As above.'		
U	no	'The rule only states that there is no relation between 3 and B. It does not state whether there is a relation between other numbers and letters.'	'Logical extension of argument above.'		
6	no	'As above.'	'As above.'		
G2,	S2				
В	yes	'To see that it is not a 3.'	'To ensure that the reverse is 3.'		
3	yes	'To ensure that it is not a B.'	'To ensure that the reverse is B.'		
U	no	'It need not prove any-thing.'	'The result might be inconclusive.'		

6	no	'It need not prove anything.'	'The result might be inconclusive.'
G2,	S 6		
В	yes	'If there is a 3 on the other side, then the statement is false.'	'If there is a 3 on the other side then the statement is true.'
3	yes .	'If there is a B on the other side then the statement is false.'	'If there is a B on the other side, then the statement is true: otherwise it is false'.
U	no	'Whatever number is on the other side will not show if statement is true or false.'	'Any number may be on the other side.'
6	no	'Any letter may be on the other side, therefore no way of knowing if statement is true.'	'If numbers are fairly ran- om, then there may be any letter on the other side, thereby giving no indica- tion unless the letter is B.'

It will be noted that the protocols differentiate between the tasks, but at a more basic level they are similar. They reveal what might be called a 'secondary matching bias': The subjects not only select the values mentioned ('matching bias'), but justify their selections in terms of the named values on the other side of the card.

Discussion

In the negative task (at least when it constitutes the first test) the justifications for choice strongly suggest logically appropriate thought processes. If such reasons had been associated with the unnegated sentences used in previous studies, they would have been interpreted as indicating 'complete insight'. There are three hypotheses which attempt to reconcile matching responses with qualitative observations incompatible with such matching: The behaviorist hypothesis, the rational hypothesis and the dual process hypothesis. (These terms are intended merely as mnemonic labels which have only an approximate reference to the hypotheses enumerated.)

The behaviorist hypothesis is that the 'reasons' are verbal responses which cannot be interpreted as revealing anything about the cause of behavior. This hypothesis is dubious because reasons and tasks are correlated. On the first task 8.3 percent (1 out of 12) of the affirmative protocols are classified as 'falsification', compared with 75 percent (9 out of 12) of the negative protocols.

The rational hypothesis is that the reasons express the cause of the choices. It follows that in the negative task the appropriate reasons indicate a genuine logical insight into the structure of the problem. But it is rather implausible to suppose that a genuine insight would suddenly vanish in the subsequent affirmative task. Furthermore, when the antecedent rather than the consequent of the conditional is negated (Evans and Lynch, 1973), the subjects tended to choose the negated value which, in this case, is logically *inappropriate*. If the effect of a negative is assumed to be the same on both the antecedent and consequent of a conditional sentence, it follows that the negated consequent in the present study does not seem to confer genuine insight into the logic of the conditional. And it follows from this that the rational hypothesis is untenable. It is, however, to be regretted that we did not include test sentences with a negated antecedent in the present study.

The dual process hypothesis postulates that performance and introspection reflect different underlying processes. It makes two fundamental assumptions:

- (1) The processes underlying the reasoning performance, e.g., matching bias, are not generally available for introspective report.
- (2) Introspective accounts of performance reflect a tendency for the subject to *construct* a justification for his own behavior consistent with his knowledge of the situation.

The process specified in (2) might be termed 'rationalization' although we do not wish to emphasize the motivational aspects of the Freudian concept. More directly relevant is the work of social psychologists who have shown a need for individuals to maintain consistency between their beliefs (Zajonc, 1968).

There are several consequences of the two assumptions. It follows from (1) that introspective protocols will be of little assistance in the construction and validation of theories concerned with predicting behavior. It follows from (1) and (2) that identification of processes underlying behavior will not be much help in predicting the nature of associated introspective protocols. The latter will depend on what the subject perceives to be the requirements of the situation, in the light of which he will *interpret his own behavior*. Thus the predominance of verification and falsification among the explicitly stated

protocols of the present experiment arises from the fact that the subject was instructed to test the truth value of an affirmative and negative sentence (respectively). He tends to interpret his choice in terms appropriate to these contexts. It also follows from the two assumptions that, instead of introspection revealing the causes of behavior, behavior is seen as one of the principal causes of the introspection. This reversal of the common sense direction of causality is reminiscent of the James-Lange theory of emotion.

But in what way does the presence of a negative elicit a justification expressed predominantly in terms of falsification? One answer would be that a negative induces the idea of falsification with respect to the terms which it negates. It does appear that the matching response is susceptible to an interaction between a specified truth value and the syntactical form of the conditional. Evans (1972a) instructed subjects to construct verifying and falsifying instances of conditional sentences in which the components were systematically negated. On the first trial, under a verifying instruction, the matching response was made to the negated antecedent in only two cases out of 48; but under a falsifying instruction, it was made in 27 cases out of 48. In other words, given a sentence in the form, 'if not (L+) then (N+)', a verifying instance tended to be constructed (correctly) as (L-) and (N+), but a falsifying instance constructed (incorrectly) as (L+) and (N+) rather than the correct (L-) and (N-). This account illuminates the 'direct falsification effect', but it leaves unexplained the 'indirect falsification effect' - the selection of the unnegated antecedent terms with appropriate falsifying reasons. But perhaps the explanation is not so deep: Falsification arises because it *fits* both the responses and the context of the task.

In its strongest form the dual process hypothesis, that response determines conscious thought, may be an oversimplification. A weaker (but more plausible) assumption is that there is a dialectical relation between them: A process of rapid continuous feedback between tendencies to respond and consciousness rather than two temporally distinct phases. This 'interpretive' view of reasoning in the selection task has been well expressed by Smalley (1974): '... reasoning is not the orderly, linear process we have sometimes imagined it to be. The kind of reasoning typified by orderly syllogistic deduction from premises to conclusion [should be] replaced by a more disorderly process in which extracted features are organized into an interpretation . . . the process is disorderly in the sense that a shifting back and forth between various pieces of information may be necessary to come to an interpretation.'

The hypothesis is formulated in its strong form so that it can be tested. It is well known that the processing of a negative normally imposes more cognitive load than the processing of an affirmative and consumes more time. Hence the dual process hypothesis and the rational hypothesis generate different predic-

tions. According to the former, the cognitive load would be imposed primarily in finding the 'reasons' (justifications); according to the latter it would be imposed primarily in determining the response. Hence the dual process hypothesis would predict that the ratio of 'justification time' to 'selection time' would be *less* for affirmative than negative sentences. The rational hypothesis would predict the converse relation. However, Smalley (personal communication) has pointed out that the semantic representation of negatives may be difficult to express as sentences regardless of which hypothesis is correct. He proposes the following simpler test. The rational hypothesis would predict that affirmatives will be faster than negatives in both 'selection time' and 'justification time'. If the prediction is not confirmed, then the alternative dual process hypothesis is corroborated. This test can be made more economical and decisive by just considering 'selection time'.

Corroborative evidence for the dual process hypothesis is to be found in previous results on the four-card problem (Wason and Johnson-Laird, 1972). These experiments suggested that subjects tended to engage in highly distinctive rationalizations to preserve an initial erroneous solution in the face of contradictory evidence. For example, they may insist that the cards are 'irreversible' or acknowledge that a card falsifies the rule but then deny its relevance. Such phenomena seem to reveal distinct thought processes which fail to interact so that conflict may remain unresolved.

One experiment (Wason, 1969b) is particularly vulnerable to reinterpretation in terms of dual processing. The solution to the problem was presented, and the task was to give reasons why it was correct. All the subjects accomplished this, and the inference was that they had been prevented from imposing their own structure on the task. But an equally plausible interpretation is that the reasons did not reflect insight but were simply constructed to fit the solution. The rival interpretations could be tested by presenting erroneous solutions as 'correct'. The dual process hypothesis would predict that 'reasons' would be found to satisfy the purported correctness of any common wrong solution. It also follows that the verbal protocols (Goodwin and Wason, 1972) collected to corroborate the Johnson-Laird and Wason (1970) model would not be interpreted as independent evidence for the stages postulated in the model. The protocols reflect the choices which define the stages.

Circumstantial evidence for the hypothesis may be found in an inductive reasoning problem (Wason, 1960, 1968b). The task was to discover a rule by generating triads of numbers (with feedback about whether they conformed to the rule). It was quite frequently found that the same hypothesis about the rule would be reformulated without awareness after the first formulation had been pronounced wrong. For example, one subject announced this rule on the

basis of confirming evidence: 'The rule is to start with a basic number, then double it and thirdly multiply it by three.' On being told it was wrong, five more confirming triads were generated, and this rule was announced: 'The rule is that the second number is double the first and two-thirds of the third.' The hypothesis, x 2x 3x, (where x = any whole number) continues to exert itself unconsciously but allows a conscious displacement to fulfill the requirements of the task.

The dual process hypothesis, however, assumes that an individual's rationalizations may be wholly appropriate when the problem lies within his competence or experience. For example, when the problem is presented in a realistic guise (e.g., Johnson-Laird, Legrenzi and Sonino Legrenzi, 1972), it is not at all difficult. The familiar content obviates the matching response because the structure of the problem is readily apparent. Indeed such a response seems to be elicited by the perplexity aroused by an abstract conditional because when a simpler logical connective (e.g., the disjunctive) is substituted, there is little evidence of matching (Wason and Johnson-Laird, 1969; Van Duyne, 1973).

The matching response may represent in miniature form a feature of intuitive thought which is so characteristic of judgments in unclear situations. In this sense it may be analogous to the 'common-element fallacy' in disjunctive concept attainment (Bruner, Goodnow and Austin, 1956, p. 168). In its abstract form the four-card problem is almost certainly not a satisfactory technique for investigating how the conditional is construed in a natural language, but it may be a potential technique for investigating intuitive and ill-defined thought processes.

Such processes have been neglected by experimental psychologists, perhaps because they are the antithesis of rational thought. And yet they are, of course, a commonplace in scientific discovery. For example, mathematicians (e.g., Poincaré) sometimes report that the solutions to their problems occur 'intuitively' and that the conscious construction of the proof is worked out after the insight. Indeed, since deduction cannot generate new knowledge and induction is philosophically discredited, some kind of intuitive process is probably the source of hypotheses. We have argued that the results of the present experiment represent a (banal) example of such processes. The fact that a trivial alteration in the task predisposes the subjects to be nominally right or wrong is beside the point, and the use of conditional sentence structure to elicit the phenomena is wholly fortuitous.

REFERENCES

- Bracewell, R. J. (1974) Interpretation factors in the four card selection task. Paper read at the Trento conference on the selection task.
- Bruner, J. S., Goodnow, J. J., and Austin, G. A. (1956) A study of thinking. London, Wiley.
- Evans, J. ST. B. T. (1972a) Interpretation and 'matching bias' in a reasoning task. Q. J. exp. Psychol., 24, 193-199.
- (1972b) On the problem of interpreting reasoning data. Cog., 1, 373—384
- and Lynch, J. S. (1973) Matching bias in the selection task. *Brit. J. Psychol.*, 64, 391-397.
- Golding E., Reich, S. S., and Wason P. C. (1974) Inter-hemispheric differences in problem solving. *Perception.*, 3, 231-235.
- Goodwin, R. Q., and Wason, P. C. (1972) Degrees of insight. *Brit. J. Psychol.*, 63, 205-212.
- Johnson-Laird, P. N., and Wason, P. C. (1970) A theoretical analysis of insight into a reasoning task. Cog. Psychol., 1, 134-138.
- Legrenzi, P., and Sonino Legrenzi, M. (1972) Reasoning and a sense of reality. Brit. J. Psychol., 63, 395-400.
- Smalley, N. S. (1974) Interpreting sentences and relating them to possible instances: An information processing approach. Paper read at the Trento conference on the selection task,
- Van Duyne, P. C. (1973) A note on Evans' matching response hypothesis. Cog., 2(2), 239-242.

- (Ed.), New horizons in psychology I. Harmondsworth, Penguin.
- (1968a) Reasoning about a rule. Q. J. exp. Psychol., 20, 273–278.
- (1968b) 'On the failure to eliminate hypotheses'... a second look. In P. C. Wason and P. N. Johnson-Laird, (Eds.), Thinking and reasoning. Harmondsworth, Penguin.
- (1969a) Regression in reasoning? Brit. J. Psychol., 60, 471–480.
- (1969b) Structural simplicity and psychological complexity: Some thoughts on a novel problem. Bull. Brit. psychol. Soc., 22, 281–284.
- -- (in press) The theory of formal operations -- a critique. In B. Geber (Ed.), Structure and development in general and social psychology: The contribution of Piaget's theory. London, Routledge.
- and Johnson-Laird, P. N. (1972) Psychology of reasoning: Structure and content. London, Batsford.
- and Golding E. (1974) The language of inconsistency. *Brit. J. Psychol.*, 65. 537-546.
- Zajonc, R. B. (1968) Cognitive theories in social psychology. In G. Lindzey and E. Aronson (Eds.), *The handbook of social psychology*, Vol. 1. London Addison-Wesley.
- (1974) Realism and linguistic complexity in reasoning. *Brit. J. Psychol.*, 65, 59-67.
- Wason, P. C. (1960) On the failure to eliminate hypotheses in a conceptual task. Q. J. exp. Psychol., 12, 129-140.
- (1966) Reasoning. In B. M. Foss

Résumé

Des résultats préalables ont montré que l'introduction de negatives dans les règles utilisées dans un problème de déduction

affectait le comportement de façon systematique indépendante de la structure logique du problème. Dans l'enquête présente on demandait aux sujets de justifier leurs réponses en raisonnant sur ces règles. Conformément aux résultats préalables, les réponses étaient dominées par les termes dans les règles sans se soucier de la négation. Les justifications variaient néanmoins quand on introduisait des négatives conformes aux conséquences logiques des

réponses. L'interpretation de ces justifications comme cause du comportement paraissait peu plausible. Il est suggéré qu'elles étaient des rationalisations où qu'il y avait au moins quelque forme de traitement double entre comportement et pensée consciente.