

Facilitating insight in a reasoning task

Emilie Matarasso Roth

Subjects were presented the selection task developed by Wason (1966) using a reduced array. The majority of subjects demonstrated partial or complete insight into the task. Performance level remained high on a transfer selection task using a standard array. The results suggest that under appropriate conditions subjects spontaneously demonstrate valid inferential processes with abstract material.

A number of researchers (Henle, 1962; Staudenmayer, 1975; Revlis, 1975) have argued that errors in reasoning result primarily from performance constraints rather than a competence deficiency. Errors in reasoning are attributed to inappropriate representation of premises or to information processing limitations. The inferential processes applied to the premises are assumed to be logically valid. While recent studies of syllogistic reasoning (Ceraso & Provitera, 1971) and propositional reasoning (Taplin & Staudenmayer, 1973; Staudenmayer, 1975) provide support for this position, studies using the selection task problem devised by Wason (1966) have consistently found evidence of fallacious reasoning that does not appear to result from misinterpretation of the premises (Wason, 1969; Wason & Johnson-Laird, 1970; Wason & Golding, 1974). The purpose of the present study is to examine whether the poor performance demonstrated on the selection task reflects a deficit in the ability to make logical inferences with abstract material or whether it can be attributed primarily to performance constraints.

In the selection task, the subject is given a conditional rule ('If P then Q ') that is supposed to apply to four cards placed in front of him with one side facing upward. For instance, the statement might be: 'If there is an A on one side then there is a 2 on the other side'. The four cards would then have an A (P), a 2 (Q), a B (\bar{P}) and a 4 (\bar{Q}) on the front. (\bar{P} and \bar{Q} stand for 'something other than P ' and 'something other than Q ' respectively.) The subject's task is to decide which of the four cards need to be turned over to determine whether the rule is true or false (given that each card has a number on one side and a letter on the other). The correct response is to choose the P and \bar{Q} cards since a \bar{Q} behind the P or a P behind the \bar{Q} would falsify the rule. Turning over the \bar{P} and the Q cards provides no information about the validity of the conditional rule.

Performance on the selection task with abstract material (e.g. 'if there is an A on one side then there is a 2 on the other') has been consistently poor (Johnson-Laird & Wason, 1970; Wason & Johnson-Laird, 1972; Wason & Golding, 1974; Smalley, 1974). The majority of subjects (over 75 per cent) choose either the P card alone or both the P and Q cards. The correct response (P, \bar{Q}) and the partially correct response (P, Q, \bar{Q}) are made infrequently (less than 10 per cent). Further, subjects demonstrate primitive thought processes (Wason, 1969; Goodwin & Wason, 1972). Instead of systematically performing a combinatorial analysis of the stimuli (i.e. considering all possible symbols that might occur on the back of each card and what implications each possible combination of symbols has for the validity of the rule), they focus on the P and Q cards and ignore the other cards (Goodwin & Wason, 1972; Wason & Evans, 1975).

Johnson-Laird & Wason (1970) proposed an information-processing model of subjects' performance on the selection task. The model classifies subjects' performance as 'no insight', 'partial insight' and 'complete insight' based on the kinds of systematic errors they make on the selection task. According to the model, 'no insight' subjects focus on the cards which have symbols on the front that match the ones in the rule (P and Q). These subjects are assumed to have a verification bias since they only consider instances that verify the rule. In contrast, subjects performing at the level of 'partial insight' and 'complete insight' consider what symbols

might occur on the back of each of the four cards; that is they perform a combinatorial analysis of the stimuli. Subjects at the level of 'partial insight' select all the cards that could either verify or falsify the rule (P , Q , \bar{Q}), while 'complete insight' subjects select only cards that potentially falsify the rule (P , \bar{Q}).

Recently, Wason and Evans (Evans & Lynch, 1973; Wason & Evans, 1975), have presented evidence that subjects responding at the level of no insight are operating under a 'matching bias' to attend primarily to cards with figures on the front that match the figures mentioned in the rule rather than a verification bias. A question of concern is whether subjects perform poorly on the selection task with abstract material because of this 'matching bias' which prevents them from realizing the necessity of performing a systematic combinatorial analysis, or whether the 'matching bias' reflects an inability to perform a combinatorial analysis with symbolic material. If it can be shown that under conditions where a matching bias is reduced, people are more likely to demonstrate superior performance on the selection task, then that would suggest that the fallacious reasoning demonstrated on the selection task results at least in part from a failure to attempt a combinatorial analysis rather than from an inability to perform such analysis successfully. Such a finding would provide support for theories of human reasoning competence that assume adults possess valid inferential processes (e.g. Inhelder & Piaget, 1958; Revlis, 1975; Staudenmayer, 1975).

The present study examines subjects' performance on the selection task under conditions where the effect of the 'matching bias' should be minimized. The subjects are presented the selection task with a reduced array that does not include the P card. Since with an affirmative rule, the tendency to focus on the P card is stronger than the tendency to focus on the Q card (Evans & Lynch, 1973), eliminating the P card should weaken the effect of the 'matching bias'. The results of a study by Lunzer *et al.* (1972) suggest that subjects presented with a reduced array are more likely to select the \bar{Q} card. However, in the study by Lunzer and his co-workers, subjects were not asked to justify their selections and a transfer task with a standard array was not presented. As a result, there is no way of knowing whether subjects who select the \bar{Q} card from a reduced array appreciate its significance. As Wason & Evans (1975) have demonstrated, subjects may select the appropriate cards and even give adequate explanations for their choices without actually possessing insight into the task. In the present study subjects presented with a reduced array are asked to justify their selections. Following the reduced selection task they are presented a standard selection task with a complete array. If subjects demonstrate improved performance under these conditions, it will provide evidence that they are capable of making valid inferences with abstract material.

Method

Design

The experiment was run as a between-subjects design with a reduced selection task condition (RST) and a control standard selection task condition (SST). Subjects were run in groups of 13 to 16 people.

Subjects in both conditions were first given a training procedure designed to make sure they interpreted the rule as a conditional. Then subjects in the RST condition were given a selection task using a reduced array containing a \bar{P} , a Q and two \bar{Q} cards, followed by a selection task using a standard array. Subjects in the SST condition were given two selection tasks using a standard array.

Materials

The stimuli were cards with one of five numbers (1, 2, 3, 4 or 5) on one side and one of four geometric shapes (square, circle, diamond or triangle) on the other. The nature of the cards was described to the subjects and examples were shown prior to the selection tasks.

Procedure

Subjects were handed booklets in which to place all responses. Everyone worked through the booklets at the same pace.

Pretraining task. Subjects were given a realistic example of an 'If... then' sentence where the converse did not hold. They were told that the sentences presented during the experiment would also not imply their converse. They were then shown both sides of four cards with the following number/figure combinations: 4/triangle, 5/diamond, 2/diamond, and 5/square. Subjects were told the statement 'If a card has a five on one side then it has a diamond on the other side' referred to both sides of all four cards. They were asked to write down whether the statement was true. After all subjects responded, they were told the statement was false and asked to write which card or cards made the statement false. Finally, they were told that only the 5/square card falsified the statement. They were asked to write why each of the other cards did not make the statement false. Almost all subjects responded appropriately to all questions in the pretraining task. Only three subjects in each condition initially chose another card in addition to or in place of the $P \bar{Q}$ card. Of these, four were able to explain why the $P \bar{Q}$ card falsified the rule.

First selection task. Four cards were placed in a stand so that subjects could only see one side of each card. The four symbols appearing on the front of the cards were printed in the booklets. Below them appeared a statement in the 'If... then' form. For the SST condition the four symbols that appeared on the front of the cards corresponded to P , Q , \bar{P} and \bar{Q} . For the RST condition the four symbols corresponded to Q , \bar{P} and two \bar{Q} (the two \bar{Q} s were represented by different symbols).

Subjects were told that the statement was supposed to apply to both sides of all four cards and that it might either be true or false. They were asked to circle in their booklets the symbols for 'all those cards and only those cards' that they would need to see both sides of in order to determine whether the statement was true or false. Subjects were cautioned that it would require some thought to come up with the right solution. They were given 2 min to make their choices. Subjects were then asked to explain for each of the four cards why they did/did not need to examine it. After subjects wrote their explanations, they were allowed a second selection of cards in case they had changed their minds. Subjects were required to give an explanation for any changes they made.

Transfer selection task. Following the first selection task, subjects in both conditions were given a standard selection task. The same procedure was followed as for the first selection task. In both conditions the four symbols presented on the cards corresponded to P , Q , \bar{P} and \bar{Q} .

After subjects wrote justifications for their second selection of cards, they were told that two and only two cards needed to be selected. They were asked to indicate which cards they were, and asked to explain any changes they made.

Following the final selection of cards for the transfer selection task, subjects were asked to write down for each of the four cards presented which figure or figures, if any, would make the statement true if it appeared on the other side of the card, and which ones would make the statement false.

Subjects

Fifty-eight undergraduate students at the University of Illinois participated in the experiment in partial fulfilment of the course requirement in an introductory psychology course. The protocols of two subjects (one in each condition) were not included in the analysis of the data because they had been familiar with the selection task prior to the experiment. The results reported below are therefore based on the protocols of 28 subjects in each condition.

Results

Selection task responses

The responses subjects made on the selection task were classified into four categories based solely on the cards selected. Following Johnson-Laird & Wason (1970) the response $P \bar{Q}$ or $\bar{Q} \bar{Q}$ on the RST) was scored as 'complete insight' and the response $P Q \bar{Q}$ (or $Q \bar{Q} \bar{Q}$ on the RST) was scored as 'partial insight'. The response $P Q \bar{P} \bar{Q}$ (or $\bar{P} Q \bar{Q} \bar{Q}$ on the RST) was scored as a 'biconditional response' since it is the correct solution to the selection task when the rule is interpreted as a biconditional. All other responses were scored as 'no insight'. Table 1 shows the number of subjects in each condition whose responses fell into each of the four categories at each point in the study.

Overall subjects in the RST condition performed better than subjects in the control condition on both the first selection task and the transfer selection task. On the first selection task, subjects in the RST condition were much more likely to select the \bar{Q} card than subjects in the SST

Table 1. Number of subjects in each condition making each response on selection task

Condition	Complete insight $P\bar{Q}$ ($\bar{Q}\bar{Q}$)	Partial insight $PQ\bar{Q}$ ($Q\bar{Q}\bar{Q}$)	Biconditional $P\bar{P}Q\bar{Q}$ ($\bar{P}Q\bar{Q}\bar{Q}$)	No insight Other
First task				
SST	4	3	1	20
RST	4	15	1	8
Revised selection				
SST	5	3	1	19
RST	6	15	0	7
Transfer task				
SST	7	3	0	18
RST	6	10	2	10
Revised selection				
SST	8	2	0	18
RST	9	8	0	11
Final selection				
SST	9	—	—	19
RST	16	—	—	12

Note: In the RST (first task) the P card did not appear. The \bar{Q} card appeared twice.

condition (see Table 1). While there were no differences between conditions in the number of 'complete insight' responses, subjects in the RST condition produced reliably more 'partial insight' responses and fewer 'no insight' responses than subjects in the SST condition, Freeman & Halton (1951; Fleishman, 1977) exact test, $P < 0.01$. In both conditions the majority of subjects who chose the \bar{Q} card referred to the possibility of a P occurring behind the \bar{Q} card in their explanations and specifically mentioned that that would falsify the rule. Of the 20 subjects in the RST condition who chose the \bar{Q} card (not including 'no insight' responses) only two failed to mention the possibility of a P occurring behind the \bar{Q} card.

Transfer task. On the transfer task, subjects on both conditions were presented with a standard array. As on the first selection task, subjects in the RST condition were more likely to choose the \bar{Q} card than subjects in the SST condition (see Table 1). Again, the major difference between the two conditions was in the number of subjects making 'partial insight' responses. Subjects in the RST condition made reliably more 'partial insight' responses than subjects in the SST condition, Freeman & Halton (1951) exact test, $P < 0.05$. The number of 'complete insight' responses was about the same in the two conditions. In both conditions the majority of subjects who selected the \bar{Q} card wrote that if a P occurred in back of the card the rule would be false. Only one of the 18 subjects in the RST condition failed to mention that a P could occur behind the \bar{Q} card.

Examination of subjects' protocols

Explanations for not selecting \bar{Q} . Across both conditions, 27 subjects initially failed to choose the \bar{Q} card on the first selection task. The explanations given by most of these (19) suggested that the \bar{Q} card was not chosen because the subjects thought it was irrelevant. Phrases such as 'did not matter', 'is irrelevant' and 'did not have P (or Q) on it' were common. A second type of

explanation given by four people was that the \bar{Q} card was not chosen because 'it was already false'. Because the explanations were often written in sentence fragments it is unclear to what the 'already false' referred. The explanations given on the transfer task for not choosing the \bar{Q} card followed a similar pattern.

Questions on selection task. After the transfer selection task subjects were asked to state for each of the cards which figures appearing on the back of the card would make the statement true and which ones would make it false. Their responses provide evidence that subjects who chose the \bar{Q} card understood its significance while those who did not choose it were unaware that the \bar{Q} card could potentially falsify the rule. Fourteen of the 17 subjects who chose $P\bar{Q}$ on the revised selection of the transfer task and all 10 subjects who chose $PQ\bar{Q}$, wrote that the statement would be false if a P occurred behind the \bar{Q} card. In contrast only six of the 29 subjects who made 'no insight' responses on the transfer task stated that a P behind a \bar{Q} would falsify the rule. The difference in the responses of 'complete insight', 'partial insight' and 'no insight' subjects is significant ($\chi^2 = 26.92$, d.f. = 2, $P < 0.001$).

Subjects' responses to the final questions also indicate that subjects responding at the level of partial insight do not think that a \bar{P} behind a Q falsifies the statement. Eight of the 10 subjects who performed at the level of partial insight on the revised selection of the transfer task did not include \bar{P} behind Q as a combination that would falsify the statement; neither did 13 of the 17 subjects who performed at the level of complete insight. In contrast, 23 of the 29 subjects who made 'no insight' responses on the revised selection of the transfer task wrote that a \bar{P} behind the Q would falsify the statement.

Final selection

At the final stage of the transfer task, subjects were asked to select two and only two cards. This manipulation was included in order to force subjects performing at the level of partial insight to consider the relative importance of the Q and \bar{Q} cards. Of the 10 subjects who chose $PQ\bar{Q}$ as their revised selection on the transfer task, five chose $P\bar{Q}$ and five chose PQ as their final selection. This is exactly what would be expected by chance. Further only one of the five subjects choosing the P and \bar{Q} cards was able to explain that the Q card was unnecessary because it could not falsify the rule. Three of the five subjects who chose the P and Q card as their final response could offer no explanation for why they chose to drop the \bar{Q} card.

Revised selection

After subjects wrote explanations for their choice of cards, they were given the opportunity to revise their selection of cards. On both the first selection task and the transfer task the pattern of responses remained approximately the same after revised selections were made (see Table 1). Writing justifications for responses in itself does not appear to facilitate a gain of insight.

Discussion

The results of the study demonstrate that subjects presented with a reduced array are more likely to recognize the significance of the \bar{Q} card than subjects presented with a complete array. While 71 per cent of the subjects in the RST condition chose the \bar{Q} card, only approximately 30 per cent of the subjects in the SST condition initially selected the \bar{Q} card. Further, the reduced array does not just increase the tendency of subjects to choose the \bar{Q} card but actually serves to increase the level of insight at which the subjects perform. Subjects in the RST condition were not choosing the \bar{Q} card arbitrarily out of confusion. Those who chose the \bar{Q} card (excluding 'no insight' responses) were generally able to give a good explanation of why the card needed to be looked at.

The fact that subjects in the RST condition were more likely to select the \bar{Q} card than subjects

in the SST condition even in the transfer task with a standard array, provides strong evidence that the reduced array serves to facilitate genuine insight into the selection task. If the improved performance with a reduced array was the result of a guessing strategy or a biasing factor, then subjects should have reverted to poor performance on the transfer task with a standard array (Wason & Evans, 1975). The fact that the majority of them did not, suggests that they had gained real insight into the selection task. The fact that the insight gained with a reduced array appears to transfer to a complete array is particularly interesting since previous efforts to induce positive transfer of insight have not been successful (Johnson-Laird *et al.*, 1972; Wason & Golding, 1974).

What is particularly impressive about the improved performance obtained with a reduced array is that it could only have occurred as a result of subjects' own analysis of the problem. Subjects presented with a reduced array are treated identically to subjects presented with a complete array. They are provided with no special prior experience, prompting or memory aids; yet they spontaneously gain the insight that the \bar{Q} card should be examined. The fact that they come to appreciate the relevance of the \bar{Q} card implies that they were guided by self-generated hypothetical information.

The explanations subjects gave for their selections provide some clues as to how the reduced array facilitates partial insight into the selection task. The fact that most subjects who failed to choose the \bar{Q} card wrote that it was 'irrelevant' or 'already false' and that they were unlikely to list a P in back of a \bar{Q} as a falsifying instance of the rule, suggests that they impulsively reject the \bar{Q} card without considering what symbols might appear in the back. These results, taken in conjunction with the results of Evans & Lynch (1973) suggest that subjects performing at the level of 'no insight' attend primarily to cards with symbols mentioned in the rule. Unlike subjects making 'no insight' responses, subjects who perform at the level of partial insight or better seem to consider what figures might occur in back of each card and to understand the implications different combination of symbols have for the rule. Almost all subjects who chose the \bar{Q} card included in their explanation some reference to a P occurring behind the \bar{Q} card. In addition most of them recognized that a P behind a \bar{Q} would falsify the statement whereas a \bar{P} behind a Q would not.

The reduced array primarily serves to shift subjects performing at the level of no insight to the level of partial insight. Since subjects at the level of no insight focus on cards mentioned in the rule while other subjects do not, it is likely that the reduced array increases the level of insight at which subjects perform by inducing subjects to attend to all the cards instead of focusing. The facilitation of insight with a reduced array may occur because of the absence of the P card; alternatively, it may be due to the fact that two \bar{Q} cards appear in the reduced array making it more salient than in a standard array. The pretraining procedure used to bias the interpretation of the rule as a conditional may also have contributed to the facilitation obtained with the reduced array. Most previous studies, with the exception of Wason & Golding (1974), did not include this type of pretraining.

The reduced array causes a dramatic increase in the number of subjects performing at the level of 'partial insight', however it is less effective in increasing the number of subjects performing at the level of complete insight compared to the control group. The majority of subjects presented with a reduced array appear to perform a combinatorial analysis successfully, however, they fail to recognize that the Q card is irrelevant to the task. Examination of protocols suggests that subjects responding at the level of partial insight select the Q card because they believe it could potentially verify the rule. These subjects appear to perceive a symmetry between the Q card and the \bar{Q} card. They consider a P on the back of a \bar{Q} to falsify the rule and a \bar{P} on the back of a \bar{Q} to be irrelevant to the rule. Similarly, they consider a P on the back of a Q to verify the rule and a \bar{P} on the back of a Q to be irrelevant to the rule. The attempt in the present study to induce subjects making 'partial insight' responses to gain

complete insight by forcing them to choose only two cards failed. The fact that subjects were as likely to drop the \bar{Q} card as the Q card suggests that subjects making partial insight responses consider verifying instances to be relevant to the evaluation of the rule as falsifying instances.

The fact that 71 per cent of the subjects performed at the level of partial insight or better on the RST, suggests that subjects perform poorly on standard selection tasks because of performance constraints rather than a competence deficiency. The standard selection task with abstract material, appears on the surface to be a simple problem. If the subject is taken in by the superficial simplicity of the task, he may impulsively attempt to solve the problem by focusing his attention on cards mentioned in the rule. In that case, the subject would fail to perform a combinatorial analysis because he does not think to do it and not because he is unable to do it (Goodwin & Wason, 1972). The selection task with a reduced array allows subjects to recognize the need for a combinatorial analysis. Once subjects are alerted to the need to perform a combinatorial analysis, they are able to do it successfully. These results are consistent with the results of recent studies of syllogistic reasoning (Ceraso & Provitera, 1971) and propositional reasoning (Staudenmayer, 1975) that suggest that the inferential processes applied by subjects are valid and that errors in reasoning result primarily from misrepresentation of premises or performance constraints.

Acknowledgements

I wish to thank Don Dulaney, Allan Fleishman and Edward Shoben for their helpful comments and suggestions. The final form of this manuscript has also profited from the comments of J. St. B. T. Evans.

References

- CERASO, J. & PROVITERA, A. (1971). Sources of error in syllogistic reasoning. *Cognitive Psychology*, 2, 400-410.
- EVANS, J. ST. B. T. & LYNCH, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology*, 64, 391-397.
- FLEISHMAN, A. I. (1977). A program for calculating the exact probability along with explorations of M by N contingency tables. *Educational and Psychological Measurement*, 37, 799-803.
- FREEMAN, G. H. & HALTON, J. H. (1951). Note of an exact treatment of contingency goodness of fit and other problems of significance. *Biometrika*, 38, 141-149.
- GOODWIN, R. Q. & WASON, P. C. (1972). Degrees of insight. *British Journal of Psychology*, 63, 205-212.
- HENLE, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-378.
- INHOLDER, B. & PIAGET, J. (1958). *The Growth of Logical Thinking*. New York: Basic Books.
- 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. & WASON, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1, 134-148.
- LUNZER, E. S., HARRISON, C. & DAVEY, M. (1972). The four card problem and the generality of formal reasoning. *Quarterly Journal of Experimental Psychology*, 24, 326-339.
- REVLIS, R. (1975). Syllogistic reasoning: Logical decisions from a complex data base. In R. J. Falmagne (ed.), *Reasoning: Representation and Process*. New Jersey: Lawrence Erlbaum.
- SMALLEY, N. S. (1974). Evaluating a role against possible instances. *British Journal of Psychology*, 65, 293-304.
- STAUDENMAYER, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (ed.), *Reasoning: Representation and Process*. New Jersey: Lawrence Erlbaum.
- TAPLIN, J. E. & STAUDENMAYER, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. *Journal of Verbal Learning and Verbal Behavior*, 12, 530-542.
- WASON, P. C. (1966). Reasoning. In B. M. Foss (ed.), *New Horizons in Psychology*. Harmondsworth: Penguin.
- WASON, P. C. (1969). Regression in reasoning? *British Journal of Psychology*, 60, 471-480.
- WASON, P. C. & EVANS, J. ST. B. T. (1974-75). 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. (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*. Cambridge, Mass.: Harvard University Press.

Received 10 September 1977; revised version received 14 March 1978

Requests for reprints should be addressed to Dr E. M. Roth, Department of Psychology, University of Illinois, Champaign, Illinois 61820, USA.