REASONING AND A SENSE OF REALITY

By P. N. JOHNSON-LAIRD, PAOLO LEGRENZI AND
MARIA SONINO LEGRENZI

University College London and Instituto Superiore di Scienze Sociali, Trento

An experiment was performed to determine whether the use of realistic materials would improve performance in a deceptive reasoning problem. The task involved selecting from a set of envelopes those which, if they were turned over, could violate a given rule. The rule concerned either a realistic relation ('if a letter is sealed, then it has a 50 lire stamp on it') or else an arbitrary relation between symbols ('if a letter has an A on one side, then it has a 3 on the other side'). Twenty-two of the 24 subjects made at least one correct answer with the realistic material but only seven of them did so with the symbolic materials. The verbal formulation of the rule was also varied but yielded only a marginal interaction with the main variable. It is argued that the critical factor is the intrinsic connexion between items rather than their specific nature.

It is a well-established fact that the content of a problem may have a significant effect upon insight into its underlying structure. Perhaps the clearest demonstration of this phenomenon in a purely deductive task is Wilkins's (1928) classic study of syllogistic inference. She discovered that problems with a familiar everyday content were generally easier than those with a purely symbolic or totally unfamiliar content: her subjects committed fewer fallacies even though the familiarity of the material provided no cue to what was, or was not, the valid conclusion. To anyone approaching thinking from a strictly formal point of view (such as the logically oriented psychologist or genetic epistemologist) this finding is both surprising and perplexing because in making a deduction the same mental operations are presumed to be carried out regardless of content. Hence, why should it be harder to execute them with one sort of material than with another? Our investigation was designed to re-examine the phenomenon and to try to answer this question.

The particular deductive problem that we chose to study was one developed by Wason (1968). The subject is presented with the four cards shown in Fig. 1, together with the following rule: *If a card has an A on one side, then it has a 3 on the other side.* He knows that each card has a letter on one side and a number on the other side; and his task is to choose just those cards which it is necessary to turn over in order to discover whether the rule is true or false.

This is an extraordinarily deceptive problem even for the most intelligent of subjects. The majority of them appreciate that it is necessary to select the 'A' card: if there was a '2' on its other side, the rule would be decisively falsified. But hardly any of them appreciate the converse need to select the '2' card. Yet, if there was an 'A' on its other side, the rule would be just as well falsified as in the first case.

There are several explanations for failure in this 'selection' task (cf. Johnson-Laird & Wason, 1970); and there are several procedures which enable the individual to gain a progressive insight into it (cf. Wason, 1969). But one simple way of inducing an initially correct selection might be to change the content of the problem so as to create a sense of reality about the task.

Imagine that you are a post-office worker sorting letters. It is your job to ensure that they conform to the following rule: *If a letter is sealed then it has a 50 lire stamp*
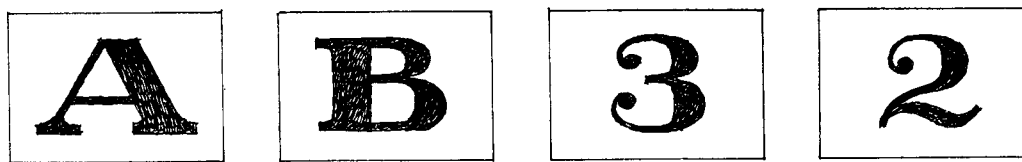
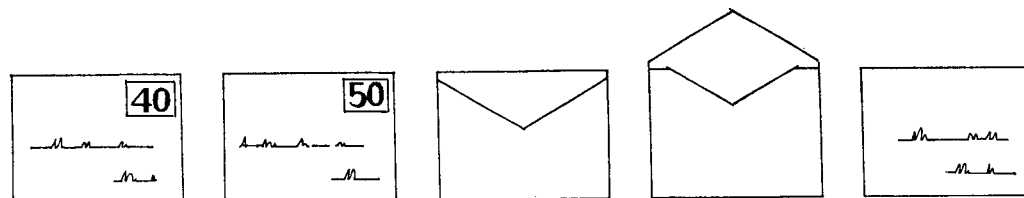Fig. 1. Four cards used in Wason's 'selection' task.

Fig. 2. The envelopes used in one condition of the experiment. The third envelope from the left was clearly sealed and the fourth envelope from the left was clearly unsealed. It was not possible to tell whether the remaining envelopes were sealed or unsealed.

on it. Which of the envelopes depicted in Fig. 2 would you need to turn over in order to discover whether or not they violated the rule? With such familiar material, where the cards themselves have become real objects from daily life, it seems relatively easy to appreciate that it is necessary to select (1) the sealed envelope (third from left in Fig. 2), (2) the envelope with the 40 lire stamp on it, and (3) the envelope with no stamp at all. The present experiment accordingly compared this 'realistic' condition with a 'symbolic' condition in which the rule referred to arbitrary numbers and letters on the envelopes. It was predicted that realism would lead the subject to a greater insight into the logical structure of the task.

The verbal formulation of the rule was also investigated, because it too seemed likely to influence insight into the task (cf. Legrenzi, 1970). Consider the following two rules:

(1) If an envelope is sealed, then it has a 50 lire stamp on it.

(2) An envelope is sealed only if it has a 50 lire stamp on it.

Logically, the two rules are equivalent: they are falsified only by a sealed envelope without a 50 lire stamp on it. There is, however, a clear difference in their meaning. Rule (1) implies that by sealing an envelope one renders it necessary to stick a 50 lire stamp on it; whereas rule (2) implies that by *not* sticking a 50 lire stamp on an envelope one renders it necessary *not* to seal it. (The difference is reflected in a more striking fashion by statements of a causal kind, e.g. 'If he's embarrassed he blushes', 'He's embarrassed only if he blushes'.) In other words, although the two formulations are identical in what renders them true or false, they differ in their implications about such contingencies. Furthermore, the formulation of rule (2) focuses attention precisely upon those contingencies customarily neglected in the selection task, i.e. upon negative instances of the second clause in the rule. This formulation should therefore produce an enhanced performance of the task. However, since this was already likely to occur with the realistic material, it was predicted that there would be an interaction between the two variables: the facilitating effect of the *only if* formulation would be greater with the symbolic material than with the realistic material.

METHOD

Design

Each subject acted as his own control and performed the four versions of the selection task generated by combining the two sorts of material (realistic *v.* symbolic) with the two sorts of verbal formulation of the rule (*if-then v. only if*). Since the lexical material was varied from one condition to another in order to minimize residual effects, the following four rules were used:

(1) If a letter is sealed, then it has a 50 lire stamp on it. (Realistic, *if-then.*)
(2) A letter is sealed only if it has a 5*d.* stamp on it. (Realistic, *only if.*)
(3) If a letter has an A on one side, then it has a 3 on the other side. (Symbolic, *if-then.*)
(4) A letter has a D on one side only if it has a 5 on the other side. (Symbolic, *only if.*)

The order of presentation of the conditions was counterbalanced, with Williams Squares, in order to detect transfer effects from one condition to another.

Materials

Five ordinary envelopes (12 × 8 cm) were used in each of the four conditions (see Fig. 2). For rule (1), they had Italian addresses, and the stamps were 40 or 50 lire. For rule (2), they had English addresses, and the stamps were 4*d.* or 5*d.* For rule (3), they consisted of the following items: 'A' on the front of an envelope, 'B' on the front of an envelope, '3' on the back of an envelope, '2' on the back of an envelope, and, finally, a blank envelope face uppermost. An equivalent set of envelopes was prepared for rule 4.

Subjects

Twenty-four undergraduates of University College London, acting as paid volunteers, were tested individually. They had no previous experience with this type of task.

Procedure

The subjects were told that the experiment concerned reasoning but that it was not a test of their intelligence. On the first occasion that a subject encountered a realistic condition he was asked to imagine that he was a post-office worker sorting letters; otherwise, he was merely told that he would have to examine some envelopes. The specific rule was then presented and the subject instructed to 'select those envelopes that you definitely need to turn over to find out whether or not they violate the rule'.

After the four selection tasks had been carried out, the subject was asked whether they were similar and, if so, in what way.

RESULTS

There was a striking difference between performance in the realistic and symbolic conditions. Table 1 shows for both conditions the number of subjects making zero, one or two correct answers. Twenty-two out of the 24 subjects produced at least one correct answer in the realistic conditions, whereas only seven of them did so in the symbolic conditions. Clearly there was considerable insight in one condition and negligible insight in the other.

The correct response, symbolizing the rules as 'if p then q' and 'p only if q', is to select the 'p' and 'not-q' envelopes. This is more insightful than the selection of 'p', 'q', and 'not-q' or of 'p', 'not-p', and 'not-q', which in turn are more insightful than the selection of 'p' or of 'p' and 'q'. Table 2 summarizes the frequencies with which such selections were made in the four conditions. In fact, 22 out of the 24 subjects made more insightful selections in the realistic conditions, and the two remaining subjects showed equal insight in the two conditions. Obviously, this is an exceedingly reliable difference ($P = 0.5^{22}$, sign test, one-tailed).

Table 1. *The numbers of subjects making a correct selection on both trials, one trial, and neither trial, in the realistic and symbolic conditions*

	Realistic conditions	Symbolic conditions
Both correct	17	0
One correct	5	7
Neither correct	2	17

Table 2. *The frequencies of the main types of selections in the four experimental conditions*

Conditions		p, not q	p, q	p	p, q, not q	Miscellaneous
Realistic	If p then q	21	1	1	1	—
	p only if q	18	—	2	1	3
Symbolic	If p then q	2	14	4	2	2
	p only if q	5	12	3	1	3

Table 3. *The distribution of the subjects in terms of whether or not they made the same selections in the two realistic conditions, and whether or not they made the same selections in the two symbolic conditions*

		Realistic conditions		
		Same selections	Different selections	
Symbolic conditions	Same selections	6	2	8
	Different selections	12	4	16
		18	6	24

There are, of course, two 'not q' envelopes in each condition: one of them has an item other than q, and the other a blank in place of q. In the majority of cases, a subject would treat both these envelopes in an identical fashion. However, on three occasions subjects selected the 'p' envelope together with just the blank envelope (the three miscellaneous selections of rule 4 in Table 2); and on another three occasions subjects selected the 'p' envelope together with just the other 'not q' envelope.

There was no overall effect of the verbal formulation of the rule upon performance. But there was a trend in the predicted direction towards a weak interaction between this variable and the content of the problem. Considering the *only if* formulation, a combination of a slight enhancement in the symbolic condition together with no detectable effect in the realistic condition, yielded results in favour of the interaction for 10 subjects and against it for four subjects, with 'ties' for the remaining 10 subjects ($P = 0.09$, sign test, one-tailed).

Finally, it was observed that the selections were more stable in the realistic than in the symbolic conditions. Table 3 gives the distribution of subjects in terms of their consistency in the two conditions. It shows that 12 subjects combined consistency in the realistic conditions with inconsistency in the symbolic conditions, whereas only two subjects behaved in the opposite fashion ($P < 0.02$, sign test, two-tailed). In the

past, selections with symbolic rules have not tended to be so labile (but cf. Wason & Johnson-Laird, 1969), hence we surmise that the present phenomenon is in part due to the effects of intermingling real and symbolic problems. However, there was a striking lack of any direct form of transfer between them; and, according to the introspective reports, only two subjects realized that there was an underlying similarity in their logical stucture.

DISCUSSION

Realism in the task induced a remarkable degree of insight into its logical structure. Indeed, the level of performance was considerably higher than any which has so far been reported for the selection problem: 91 per cent of the subjects made a correct selection in at least one of the two realistic conditions, and just over 70 per cent of the subjects made a correct selection in both of them. The only comparable level of performance was obtained by Wason & Shapiro (1971). They used a selection task with a 'thematic' rule of the form: *Every time I go to Manchester I travel by car*, together with a set of cards representing various destinations and modes of transport. The framework of this experiment is obviously imaginary because the cards which represent journeys remain simply cards. Yet it seemed to produce a sufficient sense of realism for the subjects to be able to think not of the cards themselves but of the journeys they represented, and 62·5 per cent of them succeeded in making the correct selection. In the present experiment, however, the cards are replaced by real objects, and the task simulates a feasible and realistic activity. The subject is not required to test a rule, but to test whether or not objects conform to a rule whose truth is guaranteed. But, of course, when the rule concerned an arbitrary relation between symbols on the envelopes, the subjects relapsed to the customary level of performance in the selection task. This demonstrates that the important factor *is* the realism of the task, and this is confirmed by the almost complete absence of transfer from the realistic to the symbolic conditions. It is the content of the problems which is crucial rather than their structural identity. Hence, it is scarcely surprising that the verbal formulation of the rule had at most a marginal effect upon performance.

It is by no means clear what constitutes the fundamental cause of the greater insight with the realistic materials. One simple possibility is that it is easier to appreciate the 'reversibility' of real objects. It will be recalled that subjects treat cards with 'A' on one side and '2' on the other side in different ways depending upon which side is uppermost. It seems that the reversibility of the cards temporarily escapes them: they fail to grasp that the cards are identical apart from their orientation. Working with envelopes, however, may prevent the subjects from losing this basic operational knowledge. But if this was the real explanation, a similar facilitation in performance should occur with symbolic materials. We carried out a more stringent test of this hypothesis in a pilot study. A comparison was made between symbolic material written on cards and symbolic material written on envelopes (as in the present experiment). The latter potential aid to reversibility in no way improved performance in the task.

There remain two other possible explanations. First, the individual may find many realistic objects easier to visualize, to remember, and to manipulate mentally. This is a plausible factor, but in our view unlikely to be the critical one. The materials in

400   P. N. JOHNSON-LAIRD, P. LEGRENZI AND MARIA S. LEGRENZI

many of the previous 'symbolic' experiments have been easy enough to remember or to visualize, yet the task has remained obstinately intractable. For example, Wason & Johnson-Laird (1970) used stimuli such as circles with borders round them, but only 5·5 per cent of their subjects made an initially correct selection.

Secondly, it could be the realistic *relation* between the contingencies in the rules which leads to the superior performance. The individual is used to considering the connexion between such items as destinations and modes of transport, or postal rates and envelopes. A study by Legrenzi (1971) provides circumstantial evidence for this hypothesis. He found that when subjects had to discover for themselves the nature of an arbitrary rule relating symbolic materials, their performance in a subsequent selection task was markedly superior to that of subjects who were merely given the rule to test. Discovering the rule seemed to make it almost as familiar as a realistic one. In both cases, it is presumably easier to try out the various combinations of items in a systematic fashion, and to grasp what is pertinent to the truth or falsity of the rule.

If this hypothesis is correct, then it might be possible to discover realistic connexions which in fact hinder the process of deduction. This is precisely the result obtained by Johnson-Laird & Shapiro (cf. Wason & Johnson-Laird, 1972). They constructed reasoning problems that could be solved only by making a hypothetical assumption. When the causal aspects of the premises were consistent with the logical requirements of the task, it was a simple matter to draw the correct conclusion. But when the two ran counter to one another, causal thinking prevailed and it was extremely difficult to draw the correct conclusion.

It is a task for the future to conduct direct tests of this explanation of the results, and to determine the precise nature of a realistic connexion between events. It may be that a sense of reality is no more than a feeling of familiarity.

REFERENCES

JOHNSON-LAIRD, P. N. & WASON, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychol.* **1**, 134–148.

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

LEGRENZI, P. (1971). Discovery as a means to understanding, *Q. J. exp. Psychol.* **23**, 417–422.

WASON, P. C. (1968). Reasoning about a rule. *Q. J. exp. Psychol.* **20**, 273–281.

WASON, P. C. (1969). Regression in reasoning? *Br. J. Psychol.* **60**, 471–480.

WASON, P. C. & JOHNSON-LAIRD, P. N. (1969). Proving a disjunctive rule. *Q. J. exp. Psychol.* **21**, 14–20.

WASON, P. C. & JOHNSON-LAIRD, P. N. (1970). A conflict between selecting and evaluating information in an inferential task. *Br. J. Psychol.* **61**, 509–515.

WASON, P. C. & JOHNSON-LAIRD, P. N. (1972). *Structure and Content in Reasoning: a Psychological Enquiry*. London: Batsford.

WASON, P. C. & SHAPIRO, D. (1971). Natural and contrived experience in a reasoning problem. *Q. J. exp. Psychol.* **23**, 63–71.

WILKINS, M. C. (1928). The effect of changed material on ability to do formal syllogistic reasoning. *Archs Psychol.* **16**, 1–83 (no. 102).