How diagrams can improve reasoning: Mental models and the difficult cases of disjunction and negation

Malcolm I. Bauer and P.N. Johnson-Laird

Department of Psychology
Princeton University
Green Hall
Princeton, NJ 08544
malcolm@clarity.princeton.edu
phil@clarity.princeton.edu

Abstract

We report two experiments on the effects of diagrams on reasoning. Both studies used "double disjunctions", e.g.:

Raphael is in Tacoma or Julia is in Atlanta, or both.
Julia is in Atlanta or Paul is in Philadelphia, or both.
What follows?

Subjects find it difficult to deduce a valid conclusion, such as:

Julia is in Atlanta, or both Raphael is in Tacoma and Paul is in Philadephia.

In the first study, the format of the premises was either verbal or diagrammatic, and the diagrams used icons to distinguish between inclusive and exclusive disjunctions. The results corroborated the prediction that inclusive disjunctions would be harder than exclusive disjunctions, but diagrams had no effect on performance. In Experiment 2, the diagrams made alternative possibilities explicit. The subjects responded faster and drew many more valid conclusions (nearly 30%) from the diagrams than from the verbal premises. The results support the prediction that diagrams that make explicit the alternative states of affairs needed for reasoning should improve performance. We discuss these results in terms of the theory of mental models (Johnson-Laird, 1983).

Introduction

Disjunctive reasoning pervades our lives. In a complex, technological world, we are frequently faced with many alternatives that we must reason about. Often we reason poorly, and in critical situations this can lead to disastrous results. For example, the operators of the Three Mile Island nuclear power plant assumed that the high temperature of a relief valve was due to a leak, and overlooked the possibility that the valve was stuck open. Similarly, the master of the English Channel ferry, The Herald of Free Enterprise, inferred that the bow door had been closed, and overlooked the possibility that they had been left open. Engineers at Chernobyl found a faulty explanation for the initial explosion, and failed to consider that the reactor had been destroyed. In each of these cases people erred by failing to consider alternatives. Perhaps if the systems they used made the disjunctive explanations more available, these disasters might have been averted.

In recent years, there has been much interest in how the use of visual representations might aid thinking. However, several researchers have been pessimistic about using diagrams to aid in inference, especially disjunctive reasoning. In a pioneering paper, Larkin and Simon (1987) distinguished the role of diagrams in three separate sorts of process: search, recognition, and inference. They showed how diagrams can help search and recognition, but about inference they wrote: "In view of the dramatic effects that alternative representations may produce on search and recognition processes, it may seem surprising that the differential effects on inference appear less strong. Inference is largely independent of representation if the information content of the two sets of inference rules [one operating on diagrams and the other operating on verbal statements] is equivalent---i.e. the two sets are isomorphs as they are in our examples. (p. 71)." Barwise and Etchemendy (1992) have recently argued that diagrams and pictures are good at presenting a wealth of specific conjunctive information. "It is much harder to use them," they say, "to present indefinite information, negative information, or disjunctive information." Such information is often better conveyed by sentences, and so their pedagogical program, Hyperproof, makes use of both diagrams and sentences.

Although researchers are pessimistic, existing theories do not state that diagrams will never aid in disjunctive reasoning, but rather researchers have not been able to think of how diagrams might help. At the same time, in many fields there is much interest in developing better representations for reasoning and learning (Reiser, Beekelaar, Tyle, and Merrill, 1991). The current paper gives a theoretical basis for why diagrams can help reasoning with disjunctions and negations, and describes two experiments that give empirical support to our claims.

In contrast to the pessimism expressed above, the theory of mental models (Johnson-Laird, 1983) does predict that certain sorts of diagrams should improve reasoning, and especially disjunctive reasoning. According to the theory subjects reason by a) constructing a model of a situation consistent with given information and background knowledge, b) determining a conclusion that is true of the model, and subject to certain other constraints (e.g. parsimony) and c) searching for alternative models in which the conclusion does not hold. If they cannot construct such alternative models, they assume their conclusion is true (Johnson-Laird, 1983; Johnson-Laird and Byrne, 1991). According to the theory of mental models, a major source of errors in reasoning is the difficulty people have in representing several models simultaneously. Hence, deductions in which subjects can construct a valid conclusion from only one model should be easier than those that require subjects to generalize across several models, and erroneous conclusions should be consistent with the truth of the premises, because the subjects consider some, but not all, of the possible models. These predictions, which are not made by theories based on formal rules (Braine, 1978; Rips, 1983), have been corroborated in studies of all the main sorts of deduction (see Johnson-Laird and Byrne, 1991, for a summary of the results). The difficulty in reasoning with several models simultaneously implies that any device that helps reasoners to keep track of alternative models of the premises should improve performance. *Therefore, diagrams that make explicit the alternative states of affairs needed for reasoning should improve performance.*

A phenomenon that is pertinent to our present discussion is the break-down of deductive performance with so-called "double disjunctions", such as:

Raphael is in Tacoma or Julia is in Atlanta, or both.

Julia is in Atlanta or Paul is in Philadelphia, or both.
What follows?

(cf. Johnson-Laird, Byrne, and Schaeken, 1992). Subjects find these problems exceedingly difficult compared to many other types of reasoning problems. If we can find a way to improve subjects reasoning with these problems, it is likely that the principle will apply to many other types of disjunctive reasoning. The premises for the problem above support five alternative models:

t	a	p	.
t	a		
	a	p	
	a		
t		p	

where each line represents a separate model of a possible state of affairs, "t" denotes Raphael in Tacoma, "a" denotes Julia in Atlanta, and "p" denotes Paul in Philadelphia. It is difficult to draw a valid conclusion, e.g. "Raphael is in Tacoma and Paul is in Philadephia, or Julia is in Atlanta" because one must enumerate all five models, and form a generalization that holds for them. The statement "Julia is in Atlanta" is true of the first four models, the statement "Raphael is in Tacoma and Paul is in Philadelphia" is true of the fifth model, and so a disjunction of the two: "Julia is in Atlanta, or Raphael is in Tacoma and Paul is in Philadelphia" results in a conclusion that holds across all five models. According to the model theory, inclusive disjunctions (i.e. propositions that state "or both" as above) should be harder than exclusive disjunctions (i.e. propositions that state "both not both") because more models are required. In addition, a 'contrary' problem in which one atomic expression is contrary to another:

Raphael is in Tacoma or Julia is in Atlanta or both
Julia is in Seattle or Paul is in Philadelphia or both
What follows?

should harder than an 'identical' problem in which the identical proposition occurs in both disjunctions (as in the first example above). Our aim in the present study was to test the prediction that such reasoning would be improved

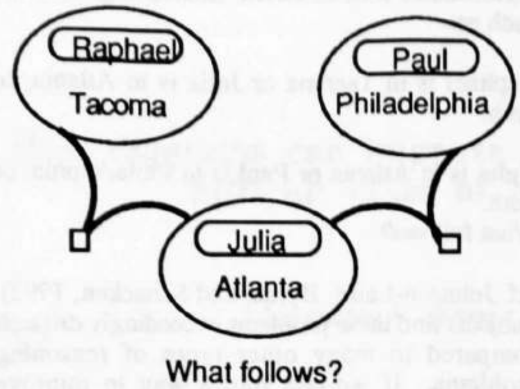


Figure 1

by the use of diagrams that make alternatives explicit. We ran two experiments to investigate this issue using the double disjunction problems. In both studies we used four double disjunction problem generated by the inclusive/exclusive and the 'identical'/contrary' dimensions. The first study used diagrams such as the one in Figure 1. The square box stands for inclusive disjunction. We used a crossed circle to represent exclusive disjunction in other problems. We ran two conditions - one consisting of the problems in diagrammatic form and the other in the verbal form. The experiment did confirm that inclusive problems are harder than exclusive, and that contrary problems are harder than identical problems. However, the diagrams did not appear to help our subjects to reason. Subjects generated 28% correct conclusions for the diagram condition and 32% correct conclusions in the verbal condition. In retrospect, the diagrams probably failed to make sufficiently explicit the alternative states of affairs, or whether the disjunctions were exclusive or inclusive. They used an arbitrary symbol to represent the form of the disjunction, e.g. a box represented inclusive disjunction. Similarly, the diagrams failed to make explicit negative instances of constituent propositions, i.e. a particular individual need not be within the oval representing a city in order to satisfy a premise. Subjects could imagine the individual outside the oval, but the diagram itself did not make this possibility explicit. In Experiment 2, we accordingly used a different sort of diagram.

Experiment 2

The aim of the experiment was to test whether subjects would reason more accurately with diagrams that made alternative and negative states of affairs explicit than with verbal premises. To examine the generality of the phenomenon, we used two different domains, one concerned people and places (as in Experiment 1) and the other switches and lights.

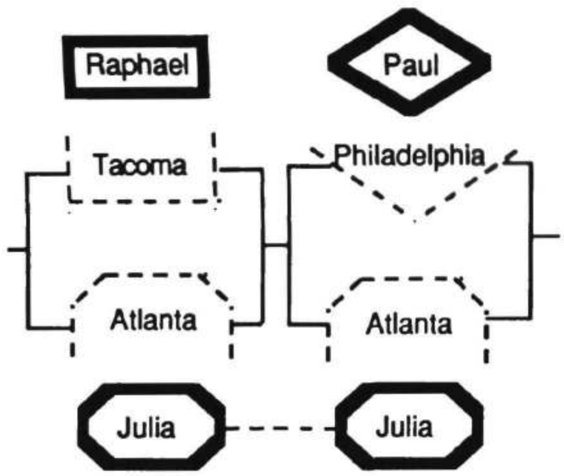
METHOD

Design: We tested four independent groups of subjects based on two factors: whether the domain was people and places or electrical circuits, and whether the format was verbal or diagrammatic. Because we found no difference between the domains for the measures relevant to this paper, we will describe only the people-and-place domain. The problems were double disjunctions of the same sort as Experiment 1: each group carried out four problems based on exclusive and inclusive disjunctions that were identical or contrary. We used Williams squares to counterbalance the order of presentation of the four problems over the subjects in each group.

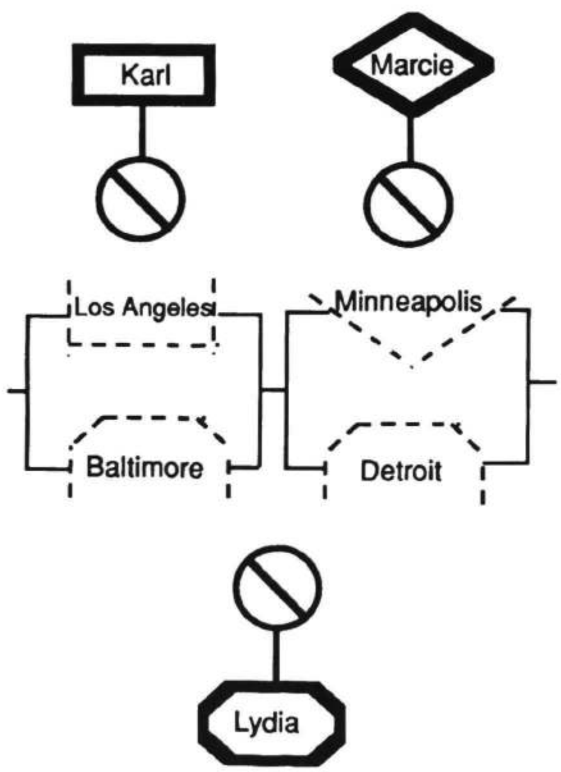
Materials: The verbal representation of the people-place domain was similar to Experiment 1. However, the diagrammatic representation of disjunction and negation made explicit the alternative possibilities. Figure 2 shows the diagram for the people-and-places problem of the identical inclusive disjunction problem, "Raphael is in Tacoma or Julia is in Atlanta or both. Julia is in Atlanta or Paul is in Philadelphia or both. What follows?" The subjects understood that they had to complete a path from one side of the figure to the other by inserting the shapes corresponding to people into the slots in the path corresponding to places. As the figure shows, the shape corresponding to a person can fit only into a similarly shaped slot corresponding to a place. Hence, in this case, the shape designating Raphael could be in Tacoma or not in Tacoma (i.e. not fitted into the Tacoma slot). The Julia shape is duplicated to form the "barbell" shape because she occurs in both premises of the verbal representation of the problem. To represent exclusive disjunction , the two disjunct shapes had attachments that occupied each other's slot, and so if one shape was inserted into its slot the other was prevented from being inserted into its slot. Figure 3 gives an example of this for a contrary exclusive problem.

Subjects: We tested 48 Princeton University undergraduate volunteers (12 in each group), and none of them had any training in formal logic or circuit analysis. They were paid $5 for participating in the experiment, which lasted for about 35 minutes.

Procedure: The subjects were tested individually. They were told that they would



The event is occurring.
What follows?
Figure 2



The event is occurring.
What follows?
Figure 3

solve a practice problem and four other problems. The task was explained to them. In the verbal conditions we explained the interpretation of inclusive and exclusive disjunctions and went on to explain that one person could not be in two places at once. In the diagram conditions subjects were given the meaningful diagram connectors and then were told how they could be interpreted, analogous to the meaning of disjunction in the verbal conditions.

Results

In contrast to experiment 1, the diagrammatic presentation increased accuracy in a striking way: the subjects drew 74% correct conclusions in comparison with 46% correct conclusions for the verbal problems ($F(1,44) = 14.9$, $p < 0.001$). The results corroborated the predicted difficulty with contrary problems (54% correct compared with identical problems 67% correct, $F(1,44) = 4.2$, $p < 0.05$)) and the greater ease of exclusive (76%) over inclusive (45%) disjunction ($F(1,44) = 24.3$, $p < 0.001$). The analyses are based on pooling the data from the circuits and people-and-places problems; this variable had no effect on either accuracy or speed of response. We examined the mean response times (in seconds) for the problems. Because there were few correct responses for the inclusive disjunction problems in the written format, we could not make a statistical test of the response times for just the correct conclusions. Instead, we analyzed the response times for all conclusions, both correct and incorrect. The subjects responded reliably faster to the diagrams (99 seconds) than to the verbal problems (135 seconds: $F(1,44) = 8.59$, $p < 0.01$). The subjects responded faster to exclusive disjunctions (104 seconds) than to inclusive disjunctives (130 seconds: $F(1,44)=24.3$, $p < .001$), and they responded faster to identical problems (97 seconds) than to contrary problems (137 seconds: $F(1,44)=52.2$, $p < .001$). In summary, the subjects were faster and more likely to make a valid inference from the diagrams than from the verbal premises, and inclusion and contrary problems were harder as predicted.

General Discussion

In Experiment 1 diagrams did not improve reasoning with double disjunctions, but in Experiment 2 they had a massive effect -- a 30% improvement. The diagrams in Experiment 1 called for considerable interpretation: disjunction was signaled only by an icon, and the difference between inclusive and exclusive disjunction by a difference in the shape of the icon. In Experiment 2, however, the diagrams graphically depicted the alternative states of affairs including the contrast between negation and affirmation. In reasoning with these diagrams, the subjects were faster and more accurate than in reasoning from verbal premises. Although the diagrams were static, the subjects could readily envisage alternative ways of completing the circuit or path by moving this piece or that into its appropriate slot (the disjunctive alternatives), and they could readily envisage the completion of one part of a path as opposed to leaving it incomplete (the contrast between affirmation and negation). It appears that diagrams that make the alternative states of affairs in a problem more explicit help subjects to reason more rapidly and accurately.

We draw two main conclusions. First, a diagram per se may not be helpful, but certain diagrams can help subjects to reason more rapidly and more accurately -- even in the case of disjunctions. Second, if diagrams are to help reasoners, they should represent alternative states of affairs explicitly. Finally, although we should like to claim to have discovered that diagrams can improve disjunctive reasoning, we note that Simon (1991, p. 96) reports anecdotally that engineers understood Supreme Court cases better when he taught in terms of switch positions in circuit diagrams representing the yes/no decisions of the court.

References

Barwise, J., and Etchemendy, J. (1992). Hyperproof: Logical reasoning with diagrams. In Narayanan, N.H. (Ed.) AAAI Spring Symposium on Reasoning with Diagrammatic Representations. March 25-27th, Stanford University, California. (pp. 80-84)

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. Psychological Review, 85, 1 - 21.

Johnson-Laird, P. N. (1983). Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., and Byrne, R. M. J. (1991). Deduction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P.N., Byrne, R.M.J., and Schaeken, W. (1992). Propositional reasoning by model. Psychological Review, 418-439.

Larkin, J., and Simon, H. (1987). Why a diagram is (sometimes) worth 10,000 words. Cognitive Science, 11, 65-99.

Reiser, B.J., Beekelaar, R., Tyle, A., and Merrill, D. C. (1991). Scaffolding learning to program with reasoning congruent representations. The International Conference of the Learning Sciences: Proceedings of the 1991 Conference. Evanston, Il.:AERA.

Rips, L.J. (1983). Cognitive processes in propositional reasoning. Psychological Review, 90, 38-71.

Simon, H.A. (1991). Models of My Life. New York: Basic Books.