Conjunctions are easier than disjunctions: A study of logical reasoning problems in the GRE

Yingrui Yang (yangyri@rpi.edu)
Department of Cognitive Sciences
Department of Decision Sciences and Engineering Systems
Rensselaer Polytechnic Institute
Carnegie Hall, Troy, NY 12180 USA

P.N. Johnson-Laird (phil@princeton.edu)
Department of Psychology
Princeton University
Green Hall, Princeton, NJ 08544 USA

Abstract

We report three experiments corroborating a prediction of the theory of mental models about reasoning. In realistic problems deriving from those used in the Graduate Record Examination (GRE), logically untrained individuals are able to cope better with conditional assertions, which have only a single explicit mental model, than with disjunctive assertions, which have multiple explicit mental models. The experiments showed that a manipulation of a sentence in the text of the problems had only a marginal effect, whereas a manipulation of the response options – whether they were both conditionals or both disjunctions – had robust effects both on the accuracy of performance and the latency of response.

Introduction

In logic, a conditional, such as: "If the trend continues then a decline will occur" is treated as equivalent to a disjunction: "Either the trend stops, i.e., doesn't continue, or a decline will occur". Both assertions are compatible with the following three possibilities:

Trend Decline
¬ Trend Decline
¬ Trend ¬ Decline

where "¬" signifies negation. In daily life, of course, the equivalence may break down as a result of the specific content or context of assertions (Johnson-Laird and Byrne, 2002). But, where the two assertions are compatible with the same possibilities, psychologists can ask: are they equally easy to comprehend and to use to make inferences? Theories based on formal rules postulate separate rules of inference for the two connectives, such as: *A or B; Not A; therefore, B*; and *If A then B; A; therefore, B*. These theories then use empirical data to estimate the relative ease of successfully using the various rules for the different connectives (see, e.g., Rips, 1984; Braine and O'Brien, 1998). In contrast, the theory of mental models predicts a difference between them (see, e.g., Johnson-Laird and Byrne, 1991). The theory postulates that individuals represent the conditional above in the mental models:

Trend Decline
. . .

The first of these models represents the possibility in which the trend continues and the decline occurs, and the second model, which has no explicit content, represents the possibilities in which the antecedent of the conditional is false. If reasoners remember that the antecedent is false in this implicit model then they can flesh it out, where necessary, into the two further possibilities shown above. The theory postulates that individuals represent the disjunction above in the mental models:

¬ Trend
Decline

And they may also add the possibility in which both propositions occur:

¬ Trend Decline

One consequence of this account is that it takes work for individuals to grasp that they can infer the conditional from the disjunction, and vice versa (see Richardson and Ormerod, 1997). Another consequence, however, is that inferences based on conditionals and that do not call for models to be fleshed out fully explicitly should generally be easier than inferences based on disjunctions. The former call for only one explicit mental model to be held in mind, whereas the latter call for multiple explicit mental models to be held in mind.

In previous research, we have investigated some of the factors that affect the difficulty of the so-called "logical reasoning" (LR) problems in the GRE developed by Educational Testing Service of Princeton, New Jersey (Yang and Johnson-Laird, 1999, 2001). The aim of the present research was to examine a further factor – the nature of the sentential connectives in the problems, and to test the prediction that problems based on conditionals would be easier to cope with than problems based on disjunctions.

Experiment 1

Method

Design The participants acted as their own controls and carried out two sorts of modified LR problems from the GRE: Problems with a conditional statement in the text and problems with a disjunctive statement in the text. Eight

original logical reasoning problems were selected from a sample of 120 representative items provided by the Educational Testing Service (ETS). Each of the 8 problems was modified to yield two versions: A conditional version and a disjunctive version. In addition, only two of the five response options from the original problems were used: One was the correct response and the other was the most frequently chosen item among the four foils (as shown in the results, which were also provided by ETS). Below are examples of the two sorts of problems used in the experiment. The two problems share the same two response options.

1. Conditional version:
Because the number of surgeons is growing faster than the number of operations and because noninvasive medical therapies are increasingly replacing surgery, the average annual number of operations per surgeon has fallen by one-fourth in recent years. It can be concluded that, if these trends continue, a dangerous decline in the level of surgical skill will occur.

2. Disjunctive version:
Because the number of surgeons is growing faster than the number of operations and because noninvasive medical therapies are increasingly replacing surgery, the average annual number of operations per surgeon has fallen by one-fourth in recent years. It can be concluded that, either these trends stop, or a dangerous decline in the level of surgical skill will occur.

The argument is based on which of the following assumptions?
(A) Surgeons now spend a large percentage of their time performing noninvasive medical procedures.
(B) A surgeon's skill cannot be properly maintained unless the surgeon performs operations with a certain minimum frequency.

Option (B) is the correct answer. By modifying the 8 original LR problems in a similar way, 16 experimental problems were obtained.

Materials. The 16 experimental problems were divided into two sets of 8 problems: Each set contained only one version of an original LR problem and included 4 conditional problems and 4 disjunctive ones. Each of these sets was printed in a booklet with one problem on each page, with enough space for the participants to make notes. The participants were given a single booklet of problems, and so they encountered a particular content only once, and carried out 4 problems of each sort.

Procedure. The participants were tested individually in a quiet room. They were given written instructions, which included one practice problem. They were encouraged to write or to draw whatever they had in mind on the problem page during the course of solving a problem. Their task was to choose whichever of the two response options they thought was correct, and they were told that there was no time limit.

Participants. Twenty undergraduates at Rensselaer Polytechnic Institute were paid to participate in the experiment. They had no training in logic.

Results and Discussion

The overall percentage of correct responses to the conditional problems was 74% and to the disjunctive problems was 67%. This difference was not significant. But, the participants were faster to solve the conditional problems (a mean of 1.76 minutes per problem) than to solve the disjunctive problems (a mean of 2.06 minutes per problem), and the difference was significant (Wilcoxon $T = 45$, $p < .05$). This result was promising, and suggested a modification to the problems, which we examined in the next experiment.

Experiment 2

Method

Experiment 2 was similar to the previous experiment, but in addition the response options for the conditional problems were both conditionals, and the response options for the disjunctive problems were both disjunctions. An example of two sorts of problem is given below:

1. Conditional version:
According to one psychological theory, if you don't have an intimate relationship with another person, then you are sad. Yet the world's greatest composers spent most of their time in solitude and had no intimate relationships. So the psychological theory must be wrong.
The argument is based on which of the following assumptions?
(A) If you are one of the world's greatest composers, then you are not sad.
(B) If you spend a lot of time in solitude then you don't have an intimate relationship.

2. Disjunctive version:
According to one psychological theory, either you have an intimate relationship with another person, or you are sad. Yet the world's greatest composers spent most of their time in solitude and had no intimate relationships. So the psychological theory must be wrong.
The argument is based on which of the following assumptions?
(A) Either you are not one of the world's greatest composers, or you are sad.
(B) Either you spend little time in solitude or you don't have an intimate relationship.

The correct answer in both cases was option A. Otherwise, the materials and procedure were identical to those of the previous experiment. Twenty undergraduates from the same population as before were paid to participate in the experiment.

Results and Discussion

The participants were more accurate with the conditional problems (73% correct) than with the disjunctive problems (61% correct, Wilcoxon's T = 13.5, $p < .005$). The participants also responded faster to the conditional problems (mean = 0.8 minutes) than to the disjunctive problems (mean = 1.17 minutes; Wilcoxon's T = 13, $p < .005$). The use of a sentence containing a given connective in both the text and the two response options evidently amplified the difference between conditionals and disjunctions. But, it leaves unanswered the question of which is the key location giving rise to the difference: the sentence in the text or the sentences in the response option. Our final experiment was designed to answer this question.

Experiment 3

Method

Experiment 3 examined all four possibilities depending on whether the sentence in the text was a conditional or a disjunction, and whether the pair of sentences in the response options were both conditionals or both disjunctions. Hence, we derived four sorts of problem from each of the original 8 LR problems, as shown here:

1. Conditional in text and conditional response options (Conditional-conditional):

Inspection system X and inspection system Y, though based on different principles, each detect all product flaws, but they each also erroneously reject three percent of flawless products. Since false rejections are very costly, money will be saved by installing both systems, instead of either one or the other, and if a product is rejected, then it must be found flawed by both systems X and Y.
The argument is based on which of the following assumptions?
(A) If system X rejects a flawless product among the three percent of such products, then system Y is not likely to reject it.
(B) If a flawed product is accepted, then the cost is less than the rejection of a flawless one.

2. Disjunction in text with conditional response options (Disjunction-conditional):

Inspection system X and inspection system Y, though based on different principles, each detect all product flaws, but they each also erroneously reject three percent of flawless products. Since false rejections are very costly, money will be saved by installing both systems, instead of either one or the other, and either a product is passed, or it found flawed by both systems.
The argument is based on which of the following assumptions?
(A) If system X rejects a flawless product among the three percent of such products, then system Y is not likely to reject it.
(B) If a flawed product is accepted, then the cost is less than the rejection of a flawless one.

3. Disjunction in text with disjunctive response options (Disjunction-disjunction):

Inspection system X and inspection system Y, though based on different principles, each detect all product flaws, but they each also erroneously reject three percent of flawless products. Since false rejections are very costly, money will be saved by installing both systems, instead of either one or the other, and either a product is passed, or it found flawed by both systems.
The argument is based on which of the following assumptions?
(A) Either system X accepts a flawless product among the three percent of such products, or system Y is not likely to reject it.
(B) Either a flawed product is rejected, or the cost is less than the rejection of a flawless one.

4. Conditional in text and disjunctive response options (Conditional-disjunction):

Inspection system X and inspection system Y, though based on different principles, each detect all product flaws, but they each also erroneously reject three percent of flawless products. Since false rejections are very costly, money will be saved by installing both systems, instead of either one or the other, and if a product is rejected, then it must be found flawed by both systems X and Y.
The argument is based on which of the following assumptions?
(A) Either system X accepts a flawless product among the three percent of such products, or system Y is not likely to reject it.
(B) Either a flawed product is rejected, or the cost is less than the rejection of a flawless one.

Problems of the sort shown in 1 and 3 are, of course, identical to those used in the previous experiment.

The resulting 32 experimental problems were divided into four sets of 8 problems, which each contained only one version of an original LR problem, and two problems of each of four sorts. Each participant carried out a single set of problems, and so encountered a particular content only once. Otherwise, the procedure was the same as in the previous experiments. Forty undergraduates from the same population as before were paid to participate in the experiment, and they were assigned by rotation to one of the four sets of problems.

Results and Discussion

Table 1 presents the percentages of correct responses and the overall mean latencies to the four sorts of problem. The results are clear-cut. What matters is the nature of the

Table 1: The percentages of correct responses and overall mean latencies (in min.) in Experiment 3.

Problem type	Accuracy (%)	Mean latency (min.)
Conditional-conditional	66	1.28
Disjunctive-conditional	70	1.36
Disjunctive-disjunctive	44	1.43
Conditional-disjunctive	50	1.43

response options. The problems with conditional response options were evaluated more accurately and more rapidly than the problems with the disjunctive response options (Wilcoxon's $T = 63.0$, $z = 3.18$, $p < .001$; Wilcoxon's $T = 262.5$, $z = 1.99$, $p < .025$, respectively). The results also replicated those of Experiment 2: the conditional-conditional problems were evaluated more accurately and more rapidly than the disjunctive-disjunctive problems (Wilcoxon's $T = 108.0$, $z = 2.56$, $p < .005$; Wilcoxon's $T = 237.0$, $z = 2.33$, $p < .01$, respectively). Finally, disjunctive-conditional problems were evaluated more accurately, though not reliably more rapidly, than conditional-disjunctive problems (Wilcoxon's $T = 47.5$, $z = 2.56$, $p < .005$).

General discussion

In accordance with the theory of mental models, when individuals reason about problems derived from those in the GRE, they cope better with those based on conditionals than with those based on disjunctions. Experiment 1 yielded only a difference in latency, which was just reliable. But the difference reflected only a single sentence in the text, which was either conditional or disjunction. Experiment 2 yielded robust effects in accuracy and latency when the difference concerned both a sentence in the text and the two response options: the three sentences were either all conditionals or all disjunctions. Experiment 3 showed that these effects were primarily a result of manipulating the response options. This phenomenon makes sense: reasoners have to keep considering the response options throughout the process of trying to solve the problems. And the correct response hinges on the complete text, not just the single sentence that we manipulated in the text of the problems. The model theory is able to elucidate certain aspects of performance with reasoning problems in the GRE. As it predicts, reasoning from conditionals is easier than reasoning from logically equivalent disjunctions.

Acknowledgments

This research was a part of the project, eWriter, which was supported by a grant from ETS and the GRE Board to S. Bringsjord, Y. Yang, P.N. Johnson-Laird, and M. Bauer.

References

Braine, M.D.S., and O'Brien, D.P., Eds. (1998) *Mental Logic*, Mahwah, NJ: Erlbaum.

Johnson-Laird, P.N., and Byrne, R.M.J. (2002) Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, in press.

Richardson, J., and Ormerod, T.C. (1997) Rephrasing between disjunctives and conditionals: Mental models and the effects of thematic content. *Quarterly Journal of Experimental Psychology*, 50A, 358-385.

Rips, L.J. (1994). *The Psychology of Proof*. Cambridge, MA: MIT Press.

Yang, Y., and Johnson-Laird, P.N. (2001). Mental models and logical reasoning problems in the GRE. *Journal of Experimental Psychology: Applied*, Vol. 7, No. 4, 308-316.

Yang, Y., and Johnson-Laird, P.N. (1999) A study of complex reasoning: The case of GRE 'logical' problems. *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society*, 767-771.