

Inference errors in deductive reasoning

LOUIS S. DICKSTEIN

Wellesley College, Wellesley, Massachusetts 02181

Recent experimental research on logical reasoning has focused on a limited number of tasks. The present study reports normative data for college students on the Inference Test, a different type of task and a marker test for the factor of logical reasoning. Six difficult items are identified, and factors that might account for the errors are discussed. These factors include over-generalization, extrapolation, addition of information, and erroneous integration of information. More general issues, such as the failure to accept the deductive task and failure to consider alternative possibilities, are noted. In the main, these tendencies are consistent with the findings of other studies of logical reasoning.

Over the past decade, there has been an upsurge of interest in experimental studies of logical reasoning (Falmagne, 1975; Revlin & Mayer, 1978; Wason & Johnson-Laird, 1972). These studies have focused primarily on three kinds of tasks: categorical syllogisms (e.g., Dickstein, 1978; Erickson, 1978; Revlin & Leirer, 1978), linear orderings (e.g., Potts, 1972, 1974; Trabasso, Riley, & Wilson, 1975), and conditional reasoning (e.g., Staudenmayer, 1975; Staudenmayer & Bourne, 1978). While these studies have yielded a number of interesting findings and have stimulated theoretical discussion, it has been increasingly recognized that information about reasoning obtained from a broader range of tasks is needed (Fraser, 1978; Revlin & Leirer, 1978). The present study was conducted to obtain normative data on reasoning with other kinds of materials.

Research on reasoning has also been conducted within the psychometric tradition. Over the years, there have been a number of factor-analytic studies of cognitive tasks (Ekstrom, French, Harman, & Dermen, 1976). Several such studies have identified a deductive reasoning factor (e.g., Adkins & Lyster, 1952; Green, Guilford, Christensen, & Comrey, 1953), which was formerly labeled syllogistic reasoning and has more recently been labeled logical reasoning (Ekstrom et al., 1976).

One of the marker tests used to measure the factor of logical reasoning is the Inference Test. This test was suggested by a similarly named test developed by Guilford (Ekstrom et al., 1976). The test consists of 20 items. Each item contains one to three sentences of information, followed by five response alternatives. The task of the subject is to decide which conclusion may be drawn from the information without making any assumptions beyond the information provided. There is only one correct conclusion for each item. Most of the items concern real-world factual material. Some of the items contain "artificial information" (Griggs, 1978), in the sense that they do not engage real-world knowledge.

While the Inference Test is regarded as a measure of

the factor logical reasoning, the items are quite different from the items typically employed in laboratory studies of categorical syllogisms. First, laboratory studies have traditionally employed the classical syllogism in which there are three terms: the subject, the predicate, and the middle term. In contrast, the items on the Inference Test are much more varied and may contain any number of terms. Second, laboratory studies have focused on the relation of set inclusion in contrast to the Inference Test, which utilizes a wide range of relations. Third, the laboratory task contains quantified propositions that may be universal or particular, affirmative or negative. In contrast, the items on the Inference Test rarely contain quantified propositions. Finally, the laboratory task typically provides subjects with four propositional conclusions relating the subject to the predicate, as well as the option that none of these may be concluded. In contrast, the response alternatives on the Inference Test do not follow any systematic format and always contain a substantive conclusion that may be derived from the given information.

In summary, laboratory studies of syllogistic reasoning have used a carefully defined task with a restricted range of materials. In contrast, the Inference Test employs a broad range of reasoning material in a much looser format. Thus, it provides an opportunity to obtain normative information about reasoning with materials that bear a much closer resemblance to everyday thinking.

METHOD

Subjects were 157 male and female undergraduates from City College of the City University of New York and from Wellesley College. There were no significant differences between males and females or between the students at the two institutions, so the data were combined. Subjects were all volunteers.

The Inference Test was administered according to the standard instructions. There are two sections to the test, with 10 items/section. Subjects were allowed 6 min for each section. Items were scored as correct, incorrect, or omitted.

RESULTS AND DISCUSSION

Overall performance on the task was quite good. The mean number correct was 16.01, the mean number incorrect was 2.56, and the mean number omitted was 1.43. Of the items answered, 86.2% were correct. There were six items for which the error rate exceeded 20%, and the remainder of this paper is concerned with a discussion of these items. The six difficult items were Items 4, 5, 8, 9, 17, and 18, with error rates of 28.6%, 29.6%, 26.2%, 46.4%, 23.2%, and 44.3%, respectively. For each of these items, there was one dominant error, which always accounted for more than half of the errors for that item. The focus of the discussion will be on these dominant errors. Each item will be presented, followed by a discussion of the dominant error for that item.¹

Item 4: "Recently, it has been discovered that it is possible to rid an area of starlings if a recording of a starling in distress is played over loud-speakers for several evenings." The dominant error reads "When starlings hear the distress calls of other starlings, they flee."

This error may be understood as an instance of overgeneralization. The subject fails to differentiate between the recording of a starling and an actual call.

Item 5: "The combined presence of carbon, water, and temperatures at which water is ordinarily a liquid are essential to life on earth." The dominant error reads "Life on the moon seems impossible since the moon has no water."

This error may also be understood as an instance of overgeneralization. The subject assumes that the elements essential to life on earth are also essential to life on the moon. The subject fails to consider the possibility that there might be different kinds of life systems.

Item 8: "Excavations at the ruins of the ancient temple called Stonehenge indicate the temple was built upon a site previously used by two other primitive civilizations. It has been estimated that the temple was constructed between 1500 and 1400 B.C." The dominant error reads "Two earlier temples had been erected on the site."

This error may be regarded as yet another instance of overgeneralization. The subject assumes that just as the site at Stonehenge was used as a temple, so previous usages must have been as temples. There is a failure to consider the possibility of different usages.

Item 9: "The salt deposits at Natrium, which are at a depth of 6,800 feet, are in the same stratum that runs under Cleveland at a depth of 1,000 feet. This stratum gradually deepens in a southeasterly direction from Cleveland." The dominant error reads "Cleveland lies to the northwest of Natrium."

This item is the most difficult item on the Inference Test and reflects an erroneous combination process. The subject is integrating information from the two sentences

and assuming that the deepening from Cleveland is toward the depths at Natrium. This assumption is reasonable but not necessary. It is quite possible for the stratum to deepen in more than one direction from Cleveland so that the relation between Cleveland and Natrium is indeterminate.

Item 17: "The world's population is increasing at a tremendous rate but the supply of water is remaining constant." The dominant error reads "In the future, few people will have enough to drink."

The error here consists of extrapolating to a future situation from a present trend. It is possible that the population trend will change. It is also possible that new sources of water will become available.

Item 18: "Skin darkness and color depend on the concentration and color of pigment particles in pigment-bearing cells." The dominant error reads "Pigment particles are not uniformly distributed in the body."

This item is the second most difficult item on the Inference Test. While the given information does suggest that cells may vary in concentration of pigment particles, the nature of such variation is unspecified. Thus, variation could be across people or it could be across different parts of the body. The error here appears to derive from subject knowledge of the world. Subjects know that skin color is not uniform across the body. This information is then combined with the given information to produce the erroneous inference.

Several general points about subject performance on this task are noteworthy. First, while the literature on syllogistic reasoning frequently documents illogical subject performance, the performance of college students on this task is quite good. More than 80% of the subjects were correct on 14 of the 20 items. Even on the six difficult items, the correct answer was always endorsed by a majority of subjects.

Second, in all instances in which an error is made, the erroneous conclusion is compatible with the given information. The inference is erroneous because it is merely possible but not necessitated by the given information. In deductive reasoning, the only legitimate conclusions are those that are compelled by the given information. Inferences that go beyond this information are too strong. This tendency toward possible but not necessary inferences parallels the performance of subjects on indeterminate categorical syllogisms. In that situation, a large number of subjects draw propositional conclusions when no definitive conclusions may be drawn (Dickstein, 1976). This type of error may be viewed as a failure to accept the logical task (Henle, 1962). Subjects may be unwilling to restrict themselves to conclusions that are implicit in the premises. Indeed, in the present task, such conclusions are often trivial.

Third, most of the errors reported in this paper involve situations in which the subject considers one plausible possibility but fails to consider other possibilities. The failure to consider other possibilities that are also consistent with the given information makes the

inference faulty. Incomplete analysis of possibilities has been noted by Dickstein (1978) and Erickson (1978) as an important factor in performance on categorical syllogisms.

The specific kinds of errors noted in this paper are also consistent with processes that have been noted in various reasoning studies. Thus, subject incorporation of knowledge about the world, as in Item 18, can be viewed as an instance of the addition of premises discussed by Henle (1962). Revlis (1975) has also discussed how information in long-term memory may be combined with information in the problem to generate an inference.

The integration of information by subjects, as in Item 9, may reflect the reasonable presupposition that if two sentences are presented together, they must be concerned with related information. Otherwise, what is the point of joining these two assertions? This kind of presupposition has been proposed by Dickstein (1978) and Henle (1978) to account for some of the errors on categorical syllogisms.

The several examples of overgeneralization reported here are consistent with the overgeneralization reported by Dawes (1964, 1966) on a memory task. Overgeneralization and extrapolation may also be viewed as further evidence of the tendency of subjects to seek "good" conceptual figures characterized by simplicity and redundancy noted by Tsal (1977).

In summary, these data provide evidence for a variety of factors that impair accuracy in deductive reasoning. These factors are consistent with error tendencies that have been reported in other studies. There is little support here for the position advanced by Revlin and Leirer (1978) that errors in deductive inference primarily reflect faulty encoding of the premises. Such a position represents an oversimplification that cannot even account for all the data on categorical syllogisms (Dickstein, 1978).

Much of the recent work on syllogistic reasoning has been devoted to the development of descriptive process models. Now that a number of processes have been identified, it would be worthwhile to focus in the future on experimental manipulations that facilitate or inhibit the various error tendencies.

REFERENCES

- ADKINS, D., & LYERLY, S. *Factor analysis of reasoning tests*. Chapel Hill: University of North Carolina Press, 1952.
- DAWES, R. Cognitive distortion. *Psychological Reports*, 1964, **14**, 443-459.
- DAWES, R. Memory and distortion of meaningful written material. *British Journal of Psychology*, 1966, **57**, 77-86.
- DICKSTEIN, L. Differential difficulty of categorical syllogisms. *Bulletin of the Psychonomic Society*, 1976, **8**, 330-332.

- DICKSTEIN, L. Error processes in syllogistic reasoning. *Memory & Cognition*, 1978, **6**, 537-543.
- EKSTROM, R., FRENCH, J., HARMAN, H., & DERMEN, D. *Manual for kit of factor-referenced cognitive tests*. Princeton, N.J.: Educational Testing Service, 1976.
- ERICKSON, J. Research on syllogistic reasoning. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- FALMAGNE, R. (Ed.). *Reasoning: Representation and process*. Hillsdale, N.J.: Erlbaum, 1975.
- FRASE, L. Inference and reading. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- GREEN, R., GUILFORD, J., CHRISTENSEN, P., & COMREY, A. A factor-analytic study of reasoning abilities. *Psychometrika*, 1953, **18**, 135-160.
- GRIGGS, R. Drawing inferences from set inclusion information given in text. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- HENLE, M. On the relation between logic and thinking. *Psychological Review*, 1962, **69**, 366-378.
- HENLE, M. Foreward. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- POTTS, G. Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 727-740.
- POTTS, G. Storing and retrieving information about ordered relationships. *Journal of Experimental Psychology*, 1974, **103**, 431-439.
- REVLIN, R., & LEIRER, V. The effects of personal biases on syllogistic reasoning: Rational decision from personalized representations. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- REVLIN, R., & MAYER, R. (Eds.). *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- REVLIS, R. Syllogistic reasoning: Logical decisions from a complex data base. In R. Falmagne (Ed.), *Reasoning: Representation and process*. Hillsdale, N.J.: Erlbaum, 1975.
- STAUDENMAYER, H. Understanding conditional reasoning with meaningful propositions. In R. Falmagne (Ed.), *Reasoning: Representation and process*. Hillsdale, N.J.: Erlbaum, 1975.
- STAUDENMAYER, H., & BOURNE, L., JR. The nature of denied propositions in the conditional sentence reasoning task: Interpretation and learning. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston Wiley, 1978.
- TRABASSO, T., RILEY, C., & WILSON, E. The representation of linear order and spatial strategies in reasoning: A developmental study. In R. Falmagne (Ed.), *Reasoning: Representation and process*. Hillsdale, N.J.: Erlbaum, 1975.
- TSAL, Y. Symmetry and transitivity assumptions about a meaningful logical relation. *Quarterly Journal of Experimental Psychology*, 1977, **29**, 677-684.
- WASON, P., & JOHNSON-LAIRD, P. *Psychology of reasoning: Structure and content*. Cambridge: Harvard University Press, 1972.

NOTE

1. These items and responses are from the *Inference Test-RL-3, Kit of Factor-Referenced Cognitive Tests*, copyright 1962, 1975, by Educational Testing Service. Reprinted by permission.

(Received for publication October 1, 1980.)