Education and the selection task

SHERRI L. JACKSON and RICHARD A. GRIGGS
University of Florida, Gainesville, Florida

The present study was concerned with the effects of education level and area of expertise on performance on the standard abstract selection task. Subjects had received bachelor's or doctoral degrees. Contrary to some recent results reported by Hoch and Tschirgi (1985), no effect of education level was found. However, there was a significant effect for area of expertise. Subjects trained in mathematics performed better than subjects from computer science, electrical engineering, and the social sciences. An explanation in terms of the mathematics subjects' greater likelihood of using a disconfirmation strategy and greater familiarity with the relevant propositional logic is offered.

Since Wason's first report on the selection task in 1966, researchers have attempted to explain why performance is so poor on this particular deductive reasoning problem. Typically less than 10% of the subjects make the correct selection. Most research has focused on the effects of problem content and other task variables, with varying results (see Evans, 1982, and Griggs, 1983, for reviews of the selection task literature). Recently, however, some attention has focused on subject variables. Hoch and Tschirgi (1985) reported data that suggests an education effect; that is, subjects with more education perform better on the task. The present study examined this factor, along with another subject variable, area of training or expertise. Because the standard abstract form of the task was used, we will describe it before we further detail the Hoch and Tschirgi research and the present study.

The selection task is based on the logic of a material conditional rule. In its standard form, the problem involves a rule of the form, "If P, then Q," and an array of four cards representing P, NOT P, Q, and NOT Q, where P and Q are usually letters and numbers. For example, the subject may be presented with four cards that display C, E, 5, and 8 and the rule, "If there is a C on one side of a card, then there is a 5 on the other side." The subject is told that each card has a letter on one side and a number on the other side and is asked to select only the card or cards that need to be turned over in order to determine whether the rule is true or false with respect to these cards.

The correct answer according to standard logic is P and NOT Q (C and 8 in the example) since these are the only cards that can possibly falsify the rule. A number other than 5 may be on the other side of the C card, and a C may be on the other side of the 8 card. The most common errors made are choosing only the P card (C in the

example) or choosing the P and Q cards (C and 5 in the example).

In stark contrast to the typically observed poor performance on the standard abstract form of the task, Hoch and Tschirgi (1985) found a 48% correct selection rate by subjects who had received master's degrees. After these subjects provided explanations of their selections, they were given the opportunity to change their answers. The correct selection rate for these selections was even higher, 52%.

In line with previous findings, subjects with high school diplomas or 2-year associate degrees and subjects with bachelor's degrees performed poorly (8% and 4%, respectively, for first choices, and 8% and 12% for second choices). Both groups performed significantly worse than the subjects with master's degrees.

Because of the atypical nature of the performance of their subjects with master's degrees, Hoch and Tschirgi ran a replication study using only subjects with master's degrees. The average correct selection rate was 41% for first choices and 55% for second choices.

In this replication study, Hoch and Tschirgi (1985) used a questionnaire to attempt to differentiate solvers from nonsolvers, but no systematic differences were found. However, Hoch and Tschirgi seemed to think that their master's level subjects differed from most previously studied subject groups. For example, they pointed out that their subjects "undoubtedly [had] well-developed analytical skills," might have been "more accustomed to thinking and reasoning at an abstract level," and were "probably similar to 'scientist' subjects in previous studies" (p. 460).

The latter point is curious since "scientists" have not fared well on the standard abstract form of the task (Griggs & Ransdell, 1986; Kern, Mirels, & Hinshaw, 1983; Mahoney & Kimper, 1976; Tolbert, 1983). For example, not one of the 10 scientists with doctoral degrees made the correct selection on the standard abstract selection task in Griggs and Ransdell's study. Thus, the performance of Hoch and Tschirgi's (1985) subjects with master's

degrees is anomalous when compared with previous findings for subjects with doctoral degrees.

To further explore the possibility of an education level effect, we examined the previous selection task literature to determine whether graduate students had ever participated. In 24 of 27 studies using the standard abstract form of the task, only undergraduates served as subjects. The percentage of subjects making the correct response ranged from 0% to 21%, with a mean of 6% correct. The three remaining studies used graduate students (Lunzer, Harrison, & Davey, 1972, Experiment 3), university graduates (Bracewell & Hidi, 1974), and possibly both undergraduates and graduates (Goodwin & Wason, 1972). The percentage of subjects making the correct response in these three studies was 10%, 8%, and 19%, respectively.

To summarize, studies with the standard abstract selection task using graduate students and scientists with doctorates as subjects do not indicate an education level effect. Overall, undergraduates, graduate students, and scientists with doctorates have all performed rather poorly. Hoch and Tschirgi's (1985) findings thus appear rather atypical and require independent replication.

In our initial attempt to replicate Hoch and Tschirgi's (1985) results, we compared subjects having master's degrees with high school graduates and found no differences in performance between these two groups on an abstract selection task (see Jackson, 1986, Experiment 1). However, questionnaire data from this experiment indicated that solvers, regardless of education level, had taken significantly more math classes than nonsolvers. This finding, along with the observation that most of Hoch and Tschirgi's subjects with master's degrees had technical (math, computer science, and engineering) degrees, led to the hypothesis that Hoch and Tschirgi's results may have been due to another subject variable, area of expertise, and not to education level.

The present experiment was designed to test this hypothesis by examining education level in conjunction with area of expertise. Including area of expertise as a factor allowed us to examine possible differences between subjects holding technical versus those holding nontechnical degrees, as well as different types of technical degrees. Four areas of expertise (social science, electrical engineering, computer science, and mathematics) and two education levels (bachelor's and doctoral) were studied.

METHOD

Subjects

Eighty students and faculty members at the University of Florida served as subjects. The subjects were classified into one of two educational levels (bachelor's or doctoral), based on the highest level of schooling completed, and one of four areas of expertise (social science [sociology or history], electrical engineering, computer science, and math), based on type of degree previously attained. There were 20 subjects within each area of expertise, 10 of which held doctorates. The subjects were contacted by phone or mail and volunteered to participate.

Materials and Procedure

The standard version of the abstract selection task used by Hoch and Tschirgi (1985) was employed; a few minor modifications were made in wording due to the fact that our subjects were run individually. The four cards were in a Plexiglas case in front of the subject and were not drawn on a page of a booklet. The rule was, "If a card has a vowel on its letter side, then it has an even number on its number side;" the four cards showing A, K, 18, and 5 were displayed along with the problem in the Plexiglas case. The order of the cards was randomized across subjects.

Because Hoch and Tschirgi (1985) presented their experimental materials in a booklet, the remaining materials in the present study were presented in a booklet. The first page explained the task and how to proceed through the booklet. The subjects were then directed to the Plexiglas case to read the problem and examine the four cards. The subjects marked their responses on the second page of the booklet.

The subjects were instructed to explain the reasons for their responses on the third page. Although the subjects were allowed to refer to the previous page to formulate their reasons, they were instructed not to change their answers for any reason. They were told that if they wanted to change their answers, they could do so beneath their explanations on the third page. The subjects were instructed that it was not necessary to change their answers, but they could if they felt they had made a mistake. The last two pages of the booklet included a questionnaire designed to examine a subject's previous exposure to factors that may have led to solving the problem (e.g., specific classes taken, previous work experience, GRE scores, etc.).

The subjects were run one at a time wherever convenient for them (mostly in their offices).

RESULTS

Two types of analyses are reported. First, the first- and second-choice selection combinations are considered. That is, the analyses of subjects' initial answers and any second-choice answers (new answers given after subjects had explained the reasons for their first-choice selections) are reported. Second, the analyses of data obtained from the questionnaire are presented.

Certain analyses required the use of multidimensional contingency tables. In these cases, the data were analyzed as in Hoch and Tschirgi (1985) by using log-linear models estimated by the method of maximum likelihood. This technique is essentially a nonparametric analysis of variance yielding chi-square values instead of F ratios.

Analyses of subjects' reasons for selecting or not selecting each card are not included, because these results were similar to those of Hoch and Tschirgi (1985), and because they are not relevant to the present discussion.

Selection Combination Analyses

The frequencies of various first- and second-choice selection combinations are shown in Tables 1 and 2, for bachelor's and doctoral level subjects, respectively.

In agreement with the results of previous studies, the most common errors were selecting either P only or P and Q. Log-linear analyses of the first-choice data in Tables 1 and 2 show that only area of expertise was significant [$\chi^2(3, N=80) = 8.0, p < .05$]. Using a standard chi-square statistic, comparisons among all areas of expertise yielded significant differences between math-

Table 1
Frequency of First- and Second-Choice Selection Combinations for Bachelor's Level Subjects and Area of Expertise

Selection	Area of Expertise			
	SS	EE	CS	MA
P ONLY	3	4	4 (3)	1
P, Q	2	2	1	2 (0)
P, NOT Q*	3	2	3 (4)	5 (7)
OTHER†	2	2	2	2

Note—SS = social science; EE = electrical engineering; CS = computer science; MA = math. If the frequency of a selection combination changed following the explanation phase, the resulting frequency is given in parentheses.   *This is the correct selection.   †The frequencies of the other possible selections were all ≤ 1.

ematics and social science subjects [χ^2(1, N=40) = 5.01, p < .05], mathematics and electrical engineering subjects [χ^2(1, N=40) = 6.67, p < .001], and mathematics and computer science subjects [χ^2(1, N=40) = 3.64, p = .05].

Thirteen subjects took the opportunity to change their answers, 6 bachelor's level subjects and 7 doctoral level subjects. Of those who changed their answers, 3 bachelor's and 5 doctoral level subjects changed to the correct answer.

Analyses of the second-choice data, including subjects who did not change their first-choice selections, led to the same findings. That is, only area of expertise had a significant effect on solution rate [χ^2(3, N=80) = 12.01, p < .01]. Specifically, mathematics subjects again performed better than subjects from the other areas (all ps < .01).

Questionnaire Analyses

The questionnaire data provided little information. This is probably due to the fact that all subjects were volunteers and did not want to take the time to think of all of their previous classes or work experiences. Also, some of the graduate students and most of the professors could remember very few, if any, classes or standardized test scores.

The only difference between solvers and nonsolvers was in the number of math classes taken [χ^2(1, N=80) = 4.7,

Table 2
Frequency of First- and Second-Choice Selection Combinations for Doctoral Level Subjects and Area of Expertise

Selection	Area of Expertise			
	SS	EE	CS	MS
P ONLY	3	1	1 (0)	1
P, Q	2 (1)	3	3 (1)	1 (0)
P, NOT Q*	2 (3)	2 (3)	3 (4)	7 (9)
P, Q, NOT Q	1	1	0 (2)	0
ALL CARDS	1	2	3	1 (0)
OTHER†	1	1 (0)	0	0

Note—SS = social science; EE = electrical engineering; CS = computer science; MA = math. If the frequency of a selection combination changed following the explanation phase, the resulting frequency is given in parentheses.   *This is the correct selection.   †The frequencies of the other possible selections were all ≤ 1.

p < .05]. Solvers had taken more math courses than had nonsolvers. Fifteen hours (about the equivalent of five classes) was used as the criterion. (A criterion was employed instead of a comparison of means for the two groups because many subjects provided only approximate totals [e.g., "more than 18 hours"].)

DISCUSSION

The results of the present study do not indicate an education level effect, but rather an area of expertise effect. Mathematics subjects did very well on the standard abstract selection task. In addition, solvers had taken more math classes than had nonsolvers. Although, at first glance, these results appear quite different from those of Hoch and Tschirgi (1985), they are actually somewhat consistent because Hoch and Tschirgi primarily used subjects with technical degrees.[1]

Although the design of the present experiment and the scarcity of relevant questionnaire data limit attempts at explaining the superiority of the mathematics subjects, the results of other studies indicate some possibilities. For example, Markowitz and Tweney (1986) found that mathematicians were in fact more likely to use disconfirmation heuristics than were science majors. Thus, mathematics subjects may be more likely to use a falsification strategy on the selection task and hence correctly solve the problem.

Why would mathematicians be more likely than scientists to use a disconfirmation strategy? It may be that scientists are trained to use both confirmation and disconfirmation strategies. Mynatt, Doherty, and Tweney (1978) found that advanced undergraduate science majors needed a firm base of inductive generalization, supported by confirmatory research, as a prerequisite to the useful implementation of a falsification strategy. Tweney, Doherty, and Mynatt (1982) reported similar results for productive research scientists.

It is also probable that mathematics subjects are more familiar with the material conditional and the other rules in propositional logic. This, too, is important. Klayman and Ha (1987) argued that on hypothesis-testing tasks, most people use a general default strategy called "positive testing" (i.e., testing cases that are expected to have the property of interest), which turns out to be effective in many natural or realistic situations. They concluded that subjects need to be familiar with the formal logic behind a problem (as mathematics subjects probably would be) to know when this heuristic is not appropriate to the task at hand. Thus, the mathematics subjects would be less likely to use positive testing and more likely to use the correct disconfirmation strategy in testing the selection task rule.

Although clearly post hoc, the above discussion provides at least a plausible explanation of the superior performance of the mathematics subjects on the selection task. Regardless, the results of the present study clearly indicate that success on the standard abstract form of the selection task seems to be the product of a particular type of education or experience, rather than the result of merely attaining a higher degree.

REFERENCES

BRACEWELL, R. J., & HIDI, S. E. (1974). The solution of an inferential problem as a function of stimulus materials. *Quarterly Journal of Experimental Psychology*, 26, 480-483.

EVANS, J. ST.B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

GOODWIN, R. Q., & WASON, P. C. (1972). Degrees of insight. *British Journal of Psychology*, 63, 205-212.

GRIGGS, R. A. (1983). The role of problem content in the selection task and THOG problem. In J. St.B. T. Evans (Ed.), *Thinking and reasoning: Psychological approaches* (pp. 16-43). London: Routledge & Kegan Paul.

GRIGGS, R. A., & RANSDELL, S. E. (1986). Scientists and the selection task. *Social Studies of Science*, 16, 319-330.

HOCH, S. J., & TSCHIRGI, J. E. (1985). Logical knowledge and cue redundancy in deductive reasoning. *Memory & Cognition*, 13, 453-462.

JACKSON, S. L. (1986). *The effects of area of expertise and level of education on performance on two reasoning tasks.* Unpublished master's thesis, University of Florida, Gainesville.

KERN, L. H., MIRELS, H. L., & HINSHAW, V. G. (1983). Scientists' understanding of propositional logic: An experimental investigation. *Social Studies of Science*, **13**, 131-146.

KLAYMAN, J., & HA, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, **94**, 211-228.

LUNZER, E. A., HARRISON, C., & DAVEY, M. (1972). The four-card problem and the generality of formal reasoning. *Quarterly Journal of Experimental Psychology*, **24**, 326-339.

MAHONEY, M. J., & KIMPER, T. P. (1976). From ethics to logic: A survey of scientists. In M. J. Mahoney (Ed.), *Scientist as subject: The psychological imperative* (pp. 187-193). Cambridge, MA: Ballinger.

MARKOWITZ, L. M., & TWENEY, R. D. (1986). *Confirmatory and disconfirmatory heuristics in mathematical reasoning.* Unpublished manuscript.

MYNATT, C. R., DOHERTY, M. E., & TWENEY, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, **30**, 395-406.

TOLBERT, C. A. (1983). *The confirmation bias as a function of level of training, hypothesis form and problem content.* Unpublished master's thesis, George Mason University, Fairfax, VA.

TWENEY, R. D., DOHERTY, M. E., & MYNATT, C. R. (1982). Rationality and disconfirmation: Further evidence. *Social Studies of Science*, **12**, 435-441.

WASON, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135-151). Harmondsworth, England: Penguin.

NOTE

1. It should be pointed out that the percentages of solvers in both the technical and nontechnical groups in the Hoch and Tschirgi (1985) study were roughly the same (S. J. Hoch, personal communication, October 16, 1985). However, the sample sizes for the nontechnical groups were very small (20 in the main experiment, only 3 in the first two replication groups from Bell Laboratories, and 7 in the University of Chicago replication sample).