DOCUMENT RESUME

ED 338 491 SE 052 316

AUTHOR Piburn, Michael D.
TITLE Teaching a Hypothesis Testing Strategy to Prospective Teachers.
PUB DATE Apr 91
NOTE 26p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (Lake Geneva, WI, April 7-10, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Development; Concept Formation; Data Analysis; Elementary Education; Higher Education; *Hypothesis Testing; Learning Processes; *Learning Strategies; *Preservice Teacher Education; Science Education; *Science Teachers

ABSTRACT

Science teachers tend to focus on schemas rather than underlying structures with greater explanatory power. One such structure is the recognition that the logical statement implication can only be falsified. After instruction with a modified version of the four-card hypothesis testing task, a grup of 48 preservice elementary teachers were significantly more able to interpret logical statements, and falsifiers were more successful than verifiers. This study demonstrates that appropriate instruction can lead preservice teachers to understand the process of hypothesis formation and falsification that is the procedure of theory testing in science. (19 references) (Author)

Teaching a Hypothesis Testing Strategy to Prospective Teachers

Michael D. Piburn
College of Education
Arizona State University
Tempe, Arizona 85287-0211

a paper presented at the 64th annual meeting of the National Association for Research in Science Teaching

Lake Geneva, Wisconsin
April 10, 1981

Running Head: HYPOTHESIS TESTING

ABSTRACT

Science teachers tend to focus on schemas rather than underlying structures with greater explanatory power. One such structure is the recognition that the logical statement of implication can only be falsified.

After instruction with a modified version of the four-card hypothesis testing task, pre-service elementary teachers were significantly more able to interpret logical statements, and falsifiers were more successful than verifiers.

This study demonstrates that appropriate instruction can lead pre-service teachers to understand the process of hypothesis formation and falsification that is the procedure of theory testing in science.

Teaching a Hypothesis Testing Strategy To Prospective Teachers

Introduction

Formal operational thought is a psychological process that mirrors the nature of scientific inquiry so closely that it is little surprise that it should be heavily emphasized by science educators. This is particularly true of the schemas, such as controlling variables or conservation principles, that provided the context within which Piaget was able to examine thought processes. However, Kuhn, Amsel, and O'Loughlin (1988) have been critical of a tendency to focus on schemas, rather than to emphasize underlying structures with greater explanatory power.

Neo-Piagetianism takes at least two forms. One attempts to interpret logical thought as a consequence of a larger explanatory system, such as, perhaps, information processing theory. The other seeks to demonstrate that elements of other psychologies are entailed by Piagetian theory. Claims of the latter sort have been made, for instance, with regard to memory, cognitive style, and spatial ability.

A candidate for consideration as a component of formal reasoning is the ability to test hypotheses, and particularly those that appear in the form of implication statements such as "if p then q" or "all p's are q." Using a particularly unusual example, Hempel (1970) addressed the proposition that "all mermaids are green", and showed that the truth of this hypothesis can only be tested by a search for mermaids that are not green. Because this insight underlies the type of hypothetico-deductive thinking that is characteristic of formal operational reasoning, (dell-Aquila, Gennaro & Picciarelli, 1985; Lawson, 1987), it is especially important that it be mastered by both students and teachers.

Moshman (1979) identified three necessary components of successful hypothesis testing: a) a correct analysis of the logical relationships presented by an implication statement, b) a realization that implications are not conclusively supported by verifying data, and c) an understanding that implications can only be falsified. Hempel (1970) suggests that our difficulty in understanding that large numbers of green mermaids could not prove the hypothesis true is the result of what he calls a psychological illusion. Moshman refers, instead, to a 'nonverification insight' and a 'falsification strategy' as requirements for such understanding.

Hypothesis testing ability is often assessed with the four-card task created by Wason and Johnson-Laird (1972). Subjects are presented with four cards showing 'A', 'D', '4', '7', and told that every card has a letter on one side and a number on the other. They are then asked which cards they would turn over to test the rule that "If a card has a vowel on one side, then it has an even number on the other side."

Because the original version of the four card task was quite abstract and difficult, Bady (1979) created an alternative version with plausible scientific content. In this task, subjects are asked to consider the hypothesis that "all turtles with diamonds on their back have green bottoms". Bady and Enyeart (1978) also conceived this as an instructional tool to teach students this difficult concept in logic. In both instances, successful completion of the task or activity requires the insight that hypotheses cannot be conclusively proven, and the ability to use a falsification strategy. Since the hypothesis can only be falsified by diamond-backed turtles with bottoms that are not green, the optimal strategy is to search for turtles with diamonds on their back to prove that they are not green or turtles that are not green to prove that they have diamonds on their back. The conclusion that turtles with diamonds and green prove nothing is very difficult for many people to accept, as is the insight that hypothetical statements of this sort cannot be proven, but only disproven.

Background

One of the principal objectives of science education is to improve the reasoning abilities of students. In consequence of this, most teacher preparation programs assume that pre-service teachers will themselves acquire such skills. Indeed, it is difficult to imagine a classroom where children become adept at critical thinking abilities that the teacher has not already mastered.

A generalized ability that appears to underlie scientific reasoning is logical thought as described by Piaget (Inhelder & Piaget, 1958). Although Piaget's proposal was that formal operational reasoning ability would be arrived at by most individuals by the age of fifteen, more recent studies, and especially those in the United States, have indicated that this may be an unduly optimistic estimate.

Studies with adolescent and college-age students indicate that only about one-third of that population are able to use formal operations (Kolodiy, 1975; Griffiths 1976). From among the remainder, about half are concrete operational and half are transitional. A study of 265 subjects in the range of age from ten to fifty years revealed that this frequency of formal operational thought does not change substantially with adulthood (Kuhn, Langer, Kohlberg & Haan, 1977). In a review of twenty-five studies of adolescent and adult subjects, King reported success rates of between 40 and 70 percent, and concluded that "a sizable proportion of the normal adult population does not reason at formal levels when tested on formal operations tasks (1986, pg. 15)."

Teachers are rarely examined as a separate population. However, Reyes (1987) used Hans Furth's An Inventory of Piaget's Developmental Tasks (IPDT) to assess the cognitive development of prospective elementary and secondary teachers. Almost half (45%) of the elementary education majors were evaluated as concrete, and only 12% were formal. Secondary education students were somewhat more successful, with 21% formal and 28% concrete. The success rate of a group of non-education students enrolled in a required speech communications course was almost identical to that of the prospective elementary school teachers. Thus it would appear that the most optimistic estimate is that prospective teachers are as advanced in terms of their cognitive development as most adolescents or adults.

Moshman and Thompson (1981) cite a number of studies indicating that most people are inclined to test hypothetical statements by seeking information that would verify them, but that the falsification strategy becomes more common in the years between the seventh grade and college. For example, dell'Aquila, Gennaro and Picciarelli (1985) found that the frequency of falsification responses rose from 2.7 to 6.4 percent between grades 6 and 12, while the percentage of verification responses dropped from 19.4 to 6.4. Using an exceptionally stringent set of criteria, including a correct logical interpretation, identification of those cases that disconfirmed the hypothesis, and identification as confirmatory only those which did not include a biconditional interpretation, Moshman (1979, pg. 107) found that "the mean numbers of implication interpretations (our of 4 possible) for subjects in grades 7, 10 and college, respectively, were .75, 1.71, and 2.42." Lawson (1983) reported even more discouraging results. No subjects from a sample of students in an two introductory college biology courses chose only the falsifying instances in either of two hypothesis testing tasks. Thus, most studies continue to support the claim by dell'Aquila, Gennaro and Picciarelli (1985, pg. 71) that "few students show a real understanding of HT (hypothesis testing) logic."

Statement of the Problem

Because of the probable link between an understanding of the nature of hypothesis testing and scientific reasoning, an attempt was made to teach the Bady "turtles" activity during an elementary science methods class for college students. The research questions were as follows:

1) If pre-service teachers are taught to use falsification strategies in hypothesis testing, will transfer occur that leads to improved performance on a test of the ability to use propositional logic;

2) What relationship, if any, exists between the ability to use a hypothesis testing strategy and formal operational reasoning; and

3) Are formal operational reasoning, the ability to use propositional logic, and achievement in a science methods class related?

Method

Subjects were 48 pre-service students seeking certification and enrolled in an elementary science methods course that met in a workshop format for one week. The core of the course was a group of 10 science centers. Additional activities included films, lectures, readings, field trips and a variety of group activities.

Propositional logic was introduced on the first day through a film about Piaget's classification schema, and instruction was given on basic propositions and truth tables. Subsequently a group activity based on Bady's revision of the Wason and Johnson-Laird four-card task was conducted (Bady & Enyeart, 1978).

The "turtles" activity asks students to consider the following statements, made by two biologists who have been studying a species of turtle that occurs only in a single remote lake. They are told that all of these turtles have either diamonds or circles on their back, and are either red or green on their bottom. They are then asked how they might evaluate the following statements by these biologists:

<u>Biologist #1</u> All turtles with diamonds on their back have green bottoms.

<u>Biologist #2</u> Biologist #1 is wrong (it is not true that all turtles with diamonds on their back have green bottoms).

The activity occurs in three phases. First, individual instances (diamonds and green, diamonds and red, circles and green, circles and red) are presented, and a discussion held with the class. Next, students are presented with a "turtle pond" (overhead projection) that contains all turtles of this species, floating either on their backs or the bottoms, and asked which turtles they should turn over in order to settle the controversy. A lively discussion generally follows in which a number of points of view are hotly contested (but unresolved by the instructor). Finally, the original instances are presented again for analysis, and students fill out a form in which they identify the turtles that they would like to see in order to evaluate the hypothesis and state why they made that choice.

In subsequent class meetings, two problems that are structurally isomorphous with the "turtles" task (Mayer, 1987, pp. 34, 202) were included on daily quizzes. These were:

1) Suppose that someone made the following assertion, "If you get bitten by a black widow spider that has a red dot on its tummy, then you will get sick." Also suppose that you have four ways of testing this assertion:

   -You get bitten by a black widow spider that has a red dot, and check to see whether or not you get sick.

   -You get bitten by a black widow spider that does not have a red dot, and check to see whether or not you get sick.

-You get sick from a black widow spider bite, and then check to see whether or not the spider has a red dot on its belly.

-You do not get sick from a black widow spider bite, and then check to see whether or not the spider has a red dot on its belly.

2) If this is room 9, then it is the fourth grade.

This is not room 9. Is it the fourth grade?

Both items were graded, returned, and discussed the following day in class.

From their responses to the second item analysis at the end of the "turtles" activity, students were assigned to one of the following groups:

III. FALSIFIERS. Correctly responded that only the diamonds and red instance was relevant, and that it falsified the hypothesis;

II. TRANSITIONAL. Correctly responded to diamonds and red, but incorrectly identified the diamonds and green instance as proving the hypothesis; and

I. VERIFIERS. Incorrectly responded that they would search for diamonds and green turtles in an attempt to verify the hypothesis.

The transfer, or 'target' task for this study was the Propositional Logic Test (PLT). This was given as a pre- and post-test (Piburn, 1989). This is a 16 item test of the ability to identify instances that are allowed or not allowed by propositional statements. It has a reported reliability, for a sample of 226 students, of .82.

The Test of Logical Thinking (TOLT) was given as a pre-test. This is a 10 item paper and pencil test of formal reasoning . Its reliability coefficient, based upon a sample of 982 students, is reported to be .85 (Tobin & Capie, 1981).

Competency tests for centers required students to solve real problems similar to those with which they had been experimenting. For example, the test for a Batteries and Bulbs center required that they light a light bulb, state whether particular electrical circuit diagrams would allow a light bulb to light, or draw the wires to complete a circuit. A Relative Position and Motion assessment provided carbon paper tracks of collisions between spheres, and required students to identify the conditions under which the tracks were made.

Grades for the course were based upon a weighted average of performance on daily quizzes (10%), class participation (10%), completion of centers (25%), competency tests on centers (25%), and a student designed center (30%).

No estimates of reliability are available either for center assessments of course grade.

Results

The first research question addresses the degree to which instruction on a falsification strategy will transfer to increased success on a measure of the ability to reason about logical propositions. An initial test of this transfer hypothesis was conducted by making a comparison of pre- and post-test scores on the Propositional Logic Test. This analysis revealed a significant increase in subjects' ability to interpret logical statements (Table 1).

----

insert table 1 about here

In order to assess the degree to which transfer from trained to target task was evenly distributed across all students, subjects were grouped as falsifiers, transitional or verifiers on the basis of their performance in the "turtles" activity, and scores on the PLT were compared across these groups by means of one-way Analysis of variance (Table 2). There were significant

----

insert table 2 about here

----

differences across the groups on both the pre-test and the post-test, with falsifiers performing better than verifiers in both comparisons. The scores of transitional subjects increased until they were not significantly different from those of falsifiers on either the pre-test or the post-test. Although their scores increased as well, verifiers continued to lag behind the other subjects on the post-test.

The second research question was whether formal thought was related to reasoning about logical propositions and hypothesis testing ability. Correlations between score on the TOLT and the PLT were 0.29 for the pretest and 0.27 for the posttest, with the latter value narrowly failing to reach significance at a probability level of 0.05. When subjects were grouped according to their performance on the "turtles" task, mean scores on the TOLT increased regularly from those of Verifiers (6.0) to Falsifiers (6.9) but the differences were not statistically significant.

The final research question was about the relationships between the two measures of reasoning ability and achievement in the science methods class. A series of analyses were conducted in order to test the hypothesis that the PLT and the TOLT were correlates of success in the class. The two achievement measures that were considered were performance on competency measures for centers and final grade in the class. A series of regression analyses were conducted in which pre-test scores on the PLT were regressed first against the dependent measure, and then the PLT posttest and TOLT scores were entered. This was done in order to eliminate initial differences that may have existed among members of the class. When scores on the PLT and TOLT variables were entered against competency measures for centers, the value of Beta for the PLT (0.27) was larger than that of the TOLT (0.16). On the other hand, when they were regressed against final grade for the course, the value of Beta for the TOLT (0.29) was larger than that for the PLT (0.20).

A principal components analysis was conducted in order to assess the relationship between intellectual and achievement

----

insert table 3 about here

----

variables. This revealed a strong initial factor that contained all of the variables under consideration, as well as a second that contained only course success (Table 3).

Conclusion

Even a brief intervention in logic can improve the quality of thinking of prospective teachers. A few hours during an elementary science methods course devoted to the nature of propositions, truth tables, and the falsification insight, led to significant increases in score on an unrelated test of the ability to use propositional logic. Given the proven association between logic, formal thought and success in this methods class, the time spent on instruction in hypothesis testing was well invested.

The development of the insight that hypotheses are not subject to verification, and the development of a falsification strategy were a crucial elements in this success. After instruction, from the entire group of 48 students, 16 (33%) fully understood these concepts, 21 (44%) had partial mastery, and only 11 (23%) continued to adopt a verification strategy. The first two groups achieved a high rate of success on the Propositional Logic Test while the third continued to lag behind. A necessary step to improving performance by the latter group is a continued intervention on hypothesis testing strategies. Only when these are correctly understood will all students achieve similar high rates of success in the interpretation of logical propositions.

Although formal reasoning and reasoning about logical propositions were shown to be related, there was no similar correspondence between the falsification insight and formal thought. These two measures are probably too different from one another for transfer to occur easily, and such a brief intervention is hardly likely to influence something as major as level of operational reasoning ability. But it is not unreasonable to believe that a continuing and more intensive course of instruction might have the desired result.

Both formal operational reasoning and the ability to interpret logical propositions were significant elements in success in a science methods class. Reasoning about logical propositions might have been somewhat more important to success at center assessments, whereas formal reasoning ability might have contributed more strongly to final grade in the course. That is not unreasonable, since course grade is a more global variable, as is formal thought. Center assessments were narrower and more process centered, and the ability to utilize logical operators in problem solving could be expected to play a stronger part in their successful completion.

Teachers who teach science must themselves be aware of the speculative nature of the enterprise. Rather than presenting science as a series of "truths" to be verified, they need to teach the logical constructs of hypothesis and falsification that constitute the normative procedures of theory testing in science. The evidence from this study is that the structures of logic and the falsification insight are closely related, and that including appropriate instruction in the pre-service training of elementary school teachers can be quite rewarding.

References

Bady, R. (1979). Students' understanding of the logic of hypothesis testing. Journal of Research in Science Teaching, 16(1), 61-65.

Bady, R. & Enyeart, M. (1978). Try and prove it: An exercise in the logic of science. The Science Teacher, 45(9), 36-37.

dell'Aquila, C., di Genaro, M. & Picciarelli, V. (1985). The logic of hypothesis testing and the control of variables formal schema: is there a link? European Journal of Science Education, 7(1), 67-72.

Griffiths, D.H. (1976). Physics teaching: Does it hinder intellectual development? American Journal of Physics, 44, 81-85.

Hempel, C. (1970). Studies in the logic of confirmation. In B. Brody, (Ed.) Readings in the philosophy of science. Englewood Cliffs, NJ: Prentice Hall.

Inhelder, B. and Piaget, J. (1958). The growth of logical thinking from childhood to adolescence. New York: Basic Books.

King, P. M. (1986). Formal reasoning in adults: A review and critique. In Mines, R. A. and Kitchener, K., Adult cognitive development: Methods and models. New York: Praeger.

Kolodiy, G. (1975). The cognitive development of high school and college science students. Journal of College Science Teaching, 5, 20-22.

Kuhn, D., Amsel, E. & O'Loughlin, M. (1988). The development of scientific thinking skills. New York: Academic Press.

Kuhn, D., Langer, J., Kohlberg, L. & Haan, C. (1977). The development of logical operations in logical and moral judgment. Genetic Psychology Monographs, 95, pp. 97-188.

Lawson, A. (1983). The effects of causality, response alternatives, and context continuity on hypothesis testing reasoning. Journal of Research in Science Teaching, 20(4), 297-310.

Lawson, A. (1987). The four-card problem resolved? Formal operational reasoning and reasoning to a contradiction. Journal of Research in Science Teaching, 24(7), 611-627.

Mayer, R. E. (1987). Educational psychology: A cognitive approach. Boston: Little, Brown.

Moshman, D. (1979). Development of formal hypothesis testing ability. Developmental Psychology, 15(2), 104-122.

Moshman, D. and Thompson, P.A. (1981). Hypothesis testing in students: Sequences, stages, and instructional strategies. Journal of Research in Science Teaching, 18(4), 341-352.

Piburn, M. (1989). Reliability and validity of the Propositional Logic Test. Educational and Psychological Measurement, 49, 667-672.

Reyes, D.J. (1987). Cognitive development of teacher candidates: An analysis. Journal of Teacher Education, March-April, pp. 18-21.

Tobin, K. & Capie, W. (1981). The development and validation of a group test of logical thinking. Educational and Psychological Measurement, 41, 413-423.

Wason, P.C. & Johnson-Laird, P.N. (1972). Psychology of reasoning: Structure and content. Cambridge: Harvard University Press.

Table 1. Comparison of pre- and post-test scores on the Propositional Logic Test

	PRETEST	POSTTEST
n	48	48
mean	12.17	13.29
S.D.	3.74	3.00
t-value	2.89	
probability	.0030	

Table 2. Comparison of pre- and post-test means on Propositional Logic Test by categories of performance on "turtles" task

	PRETEST	POSTTEST
Verifiers	9.75	11.63
Intermediate	12.90	14.05
Falsifiers	14.27	14.27
F-ratio	6.87	4.25
significance	.0028	.0199

Table 3. Principle components analysis of three measures of reasoning ability, and two components of success in methods course.

Variable	Factor 1	Factor 2
Centers	0.64	0.62
Course	0.72	0.56
PrePLT	0.69	-0.60
PostPLT	0.74	-0.50
TOLT	0.59	-0.02
Eigenvalue	2.30	1.30