DOCUMENT RESUME

ED 218 591 CS 006 735

AUTHOR Flippo, Rona F.; Borthwick, Paul
TITLE Should Testwiseness Curriculum Be a Part of Undergraduate Teacher Education?
PUB DATE Dec 81
NOTE 18p.; Paper presented at the Annual Meeting of the American Reading Forum (2nd, Sarasota, FL, December 10-12, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Curriculum Development; *Curriculum Research; Educational Research; Elementary School Students; *Teacher Education; *Teacher Education Curriculum; *Test Coaching; *Test Wiseness

ABSTRACT

Although research on testwiseness has varied to a great extent in the aspects of testwiseness measured and the strategies used, some conclusions can nevertheless be drawn: testwiseness is a bona fide entity and deserves attention, it is measurable, it appears to be made up of a set of specific skills rather than a collection of some general holistic abilities, and training in testwiseness can be administered to students of all ages from early childhood to adulthood. Working on the assumption that teachers at all levels need to be made aware of testwiseness strategies and be able to train their students in acquiring these skills, a study examined whether some training of undergraduate student teachers to teach testwiseness strategies would result in a significant degree of difference between the criterion referenced content area test scores of elementary school children taught these strategies and the test scores of children not taught the strategies. Eight student teachers and 174 elementary school students participated in the study. Each day for 3 weeks students in the experimental groups received testwiseness activities while the students in the control groups worked on other activities such as art or library work. Results indicated that although there was not a significant difference between the experimental and control treatment groups for any of the units taught and for any of the grades or ability groups, there was a slightly higher observed score for six of the nine experimental groups. This does not indicate that the student teachers were able to transfer the testwiseness strategies instruction they received to the students they taught. (HOD)

SHOULD TESTWISENESS CURRICULUM BE A PART OF UNDERGRADUATE TEACHER EDUCATION?

Rona F. Flippo, Ed.D.
Georgia Department of Education

Paul Borthwick, Ph.D.
Delta State University

The idea of teaching testwiseness to school-aged students and post-secondary students is certainly not innovative or earthshaking. Interest in this area has gained momentum over the last few years. In most situations the subjects of this interest have been the students. We are seeing more books and more chapters within reading and study skills books that have been written to teach students testwiseness strategies. Although some material has been directed toward the elementary child, more efforts have centered on the high school and college student.

While we concur with these efforts and realize the need for good testwiseness books for students, we also see another need. Perhaps some effort should be made in undergraduate teacher education? If prospective teachers took testwiseness as part of their undergraduate curriculum, perhaps they could successfully transpose some of these testwiseness strategies to the students they will teach? We feel that this undergraduate training should be explored for elementary majors as well as for the secondary majors. Perhaps some early testwiseness training in the elementary grades across the disciplines would cut down on some of the potential test-taking problems in secondary schools and colleges? We know that many education programs include a class in testing and measurement. Perhaps testwiseness strategies curriculum could be a unit of such a class?

What is Testwiseness?

The term "testwiseness" (TW) has meant many things to many people. A variety of interpretations and definitions has been associated with it. The following definition of testwiseness by Millman, Bishop, and Ebel (1965) will be used for purposes of this paper:

> "A subject's capacity to utilize the characteristics and formats of the test and/or the test taking situation to receive a high score." (p. 707)

This definition implies that TW is independent from the test-takers' knowledge of the content contained on the test itself. An immediate question that comes to mind would be whether or not one's mental or general ability has a relationship with TW. A review of the literature reveals that numerous studies have investigated this question.

Diamond and Evans (1972) studied the TW of selected sixth grade students and reported that no relationship exists between TW and general skill or ability. However, they did conclude that TW was unique to the cues or clues under study. This conclusion had also been previously substantiated by Dunn and Goldstein (1959) who found no relationship between TW and intelligence.

In a related study, Ayrer, Diamond, Fishman and Green (1976) did not find any relationship existing between TW and increased scores on the California Achievement Test among inner city students. For the most part, all correlations were low. Thus, this study does not support the Diamond and Evans research statement that "TW as a secondary cue response is quite specific to the particular clue or cue under investigation."

The question as to whether TW ability of students changes as they progress through the grades was investigated by Crehan (1977). This longitudinal study indicated that TW apparently is rather stable but increases

somewhat up through the grades. These investigators also noted a lack of evidence supporting a relationship between TW and the sex of the students.

Testwiseness Strategies

When discussing TW, many references are made to the categorization of TW strategies as developed by Millman, Bishop and Ebel (1965). In their paper, two categories are presented: elements that are independent of the test constructor or test purpose and those that are dependent upon the constructor or purpose. The first category, often called general test-taking strategies, contains strategies for utilizing time efficiently, proofreading answers in order to avoid errors, learning how and when to make appropriate guesses, and using deductive reasoning in order to arrive at a correct answer. Included in the second category are strategies designed to consider the intent of the test constructor for using certain questions and for using cues found in the test when the answer is uncertain.

Four basic TW strategies were developed by Crehan (1977). Two of these strategies involve deductive reasoning and the other two strategies involve using cues in order to arrive at an appropriate answer.

Gibb (1964) developed a strategy concerned with the problem of cues in test items and identified TW as a secondary cue response. His major premise was that students can be trained in TW and their performance would not be determined by knowledge of the subject matter.

The study by Diamond and Evans (1972) used fictitious material to eliminate content familiarity with naive sixth grade students, and investigated the correlates of TW. Diamond and Evans' strategy included the following five item faults selected for the study: association, specific determiners, correct alternatives, grammatical clues and overlapping distractors. With TW defined as a secondary cue response, the naive students appeared to have possessed it. However, data revealed that TW is not a general trait, but specific to the particular clue or cue being investigated. Therefore, the results were supportive of those of both Gibb (1964) and Dunn and Goldstein (1959).

The research has revealed that TW is not a pervasive skill since consistent individual differences are found in response to a specific form of strategy clue (i.e., grammatical clues) and that these responses do not relate to one's general mental ability.

Testwiseness Training

The practice of TW training has been widely investigated with the effects resulting in varying degrees of success. Giving emphasis on characteristics and format of the test rather than upon content or coaching for a specific exam, Wahlstrom and Boersma (1968) concluded that ninth grade students receiving TW training had higher scores on criterion referenced tests. Oakland (1972) showed that disadvantaged preschool students having had TW training obtained a higher mean on the Metropolitan Reading Readiness Test than those in the control group. However, when measured four months later, the results were not significantly different.

Callenbach's (1971) investigation reported that test-naive second graders who received TW training scored significantly higher on both an immediate and delayed standardized reading posttest than those students who received no training.

Not all researchers, however, found TW training to have positive effects. As cited earlier, Diamond et al. (1976) found no substantial relationship between the scores on the Vocabulary, Reading Comprehension or Language Usage subparts of the California Achievement Test of students given TW training and the control group. Yearby (1975) found that the TW skills of some third grade students were significantly increased by training but that this did not transfer to a significant degree on a standardized reading test. In addition, although Jongsma and Warshauer (1975) report that fifth grade inner city students who had studied a unit on TW skills averaged higher reading achievement scores than a control group, their gains were not significant.

Although research on TW has varied to a great extent in the aspects of measuring TW and the strategies used, some conclusions can nevertheless be drawn. First, TW is a bona fide entity and deserves attention. Second, TW is measurable. Third, TW appears to be made up of a set of specific skills rather than a collection of some general or wholistic abilities; and, fourth, TW training can be administered to students of all ages from early childhood to adulthood.

It would, therefore, seem that teachers at all levels need to be made aware of TW strategies and be able to train their students in acquiring these skills. This awareness could take place in teacher education programs, where preservice teachers could be trained to teach TW skills to their future students.

Pilot Study

Since we believe that some emphasis on the teacher as the subject of TW interest is called for, we developed this pilot study to begin looking at that area of teacher training. Because the need for the training of secondary education majors in TW is perhaps more obvious, we decided to do this pilot with elementary education majors. We feel that this pilot study might help give us some sense of the effectiveness of this training for preservice elementary teachers who will teach in a variety of grade and ability levels and in a variety of content areas.

The purpose of this investigation was to explore the research question: Will some training of undergraduate student teachers to teach TW strategies result in a significant degree of difference between the criterion referenced content area test scores of elementary children taught these TW strategies and the test scores of elementary children not taught these TW strategies?

Procedures

Just prior to their student teaching, and while completing a testing and measurement class, eleven elementary undergraduate majors received a module (Flippo) containing information and learning activities in TW strategies for objective tests appropriate for use with early childhood through college-level students, and a seminar class conducted by one of the researchers. This class covered an orientation to the use of the module and an explanation of the procedures to be used with the module, activities for teaching TW strategies to children, and information regarding the pilot study. Of the eleven student teachers receiving the TW training, eight of them have been included in this pilot.

Student teachers were assigned to second through sixth grade ability grouped classrooms in one of four elementary schools in northwest Mississippi. Students in these schools had been previously ability grouped by their reading scores on the California Achievement Test. In most cases, student teachers were assigned to classrooms with all high ability grouped children in a given grade in the same room or all average ability grouped children in a given grade in the same room. One student teacher had a more heterogeneous classroom assignment; however, there were not enough low ability children in that room to use low ability groups in this pilot study.

Students in the high ability and/or average ability groups in each class were randomly assigned at each of these ability levels into control and experimental groups. A total of 174 randomly assigned students in grades two through six were included in this study. Seventy-one of them were in experimental groups and one hundred and three of them were in control groups. The student teachers taught a unit in one of the social studies disciplines to the entire class of children over a period of about one month. Additionally, experimental groups were given TW instruction by their student teachers. The students in the control groups did not receive TW strategies training but instead were involved in regularly planned instructional activities within the units being studied.

Each day for a little over three weeks the student teachers worked through the 16 activities in their modules with the experimental groups of children. While the experimental groups worked on the TW activities with the student teachers in the classrooms, the children in the control groups worked on other activities like art, library work, etc., related to their social studies unit with the regular classroom teachers outside the classrooms. During this same time period, student teachers conducted

whole class instruction in social studies units. (Refer to Table 2 for the units taught by the student teachers.) The student teachers were cautioned not to include any TW information when teaching the social studies units. Likewise they were cautioned not to include any material from the social studies units in the TW activities with the experimental groups.

Each student teacher developed a criterion referenced test covering the social studies unit s/he taught for the purpose of measuring students' knowledge of the content taught in that unit. This test was developed according to the guidelines and instructions that the student teacher received in her/his testing and measurement class. It also included some not so obvious TW cues. (The idea was that the testwise student would pick up on those cues and do better on the test while the student who was not testwise would not do as well.) All tests were submitted to the Division of Curriculum and Instruction at Delta State University and the researchers for evaluation of validity.

At the completion of the social studies units and treatment activities, each entire class in this study was given the social studies unit test developed by their student teacher. Only data from the tests which met a minimum quality of measurement were included in this study.

Statistical Analysis

An initial concern was the possible existence of significant disparity in TW between the experimental and control groups in each classroom prior to administration of the treatment. To check for pre-treatment variance, the classroom teachers of these groups were asked to rate the level of TW of each child in their class on a thirty point scale. The "Testwiseness Rating Scale"[1] was used. (See the scale in Figure 1, and see Table 1 for directions for its use.) Group ratings were then compared using a Mann-Whitney Test (Ryan, Joiner, & Ryan, 1976).

Of nine group comparisons made from the eight classrooms included in this study, none were significant at the .05 alpha level. Table 2 indicates median scores for the different groups and the subsequent p levels derived from administration of the Mann-Whitney Test. The findings indicated no significant difference in TW between experimental and control groups for any class prior to treatment.

(Insert Table 2 about here)

To determine effectiveness of the treatment, the unit posttest scores for each instructional group were compared. The Mann-Whitney Test was used. Table 3 indicates median scores for the different groups and the subsequent p levels. The findings indicate that the TW instruction was not a significant influence on unit posttest performance.

(Insert Table 3 about here)

Conclusions and Recommendations

Although there was not a significant difference between the experimental and control treatment groups for any of the units and for any of the grades and ability groups included in this pilot study, there was a slightly higher observed score for six of the nine experimental groups. This observed, but insignificant, difference certainly does not indicate that the student teachers were able to transfer the TW strategies instruction they received to the students they taught. However, it does indicate that a more tightly controlled study should be done to further explore the possibilities of TW training for student teachers and its transfer to their future students.

Many factors should be considered, such as the student teachers' mastery of the TW instruction they receive, the effects of several different approaches to the TW training of student teachers, the amount of TW training

given to student teachers, and the effects of providing TW training to student teachers at different times or stages in their undergraduate education preparation.

While this pilot study did give us some sense of the effectiveness of TW training for preservice elementary teachers, it suffered because of the number of concomitant variables that were not controlled for. For instance, to what extent did the outside-the-classroom enrichment activities (i.e., art, library work, etc.) in the social studies units provide the control groups with more knowledge of or interest in the units, as compared to the benefits for the experimental groups receiving TW training? Or, to what extent did the newness or novelty of the student teachers teaching the TW effect the experimental groups' attentiveness? Did all of the student teachers give equal attention to the TW instruction of their experimental groups? These are but a few of the variables that should be controlled before another similar study is attempted. We think some TW training for preservice education students is a good idea; however, more positive documentation will be necessary before it can be thoroughly recommended.

Reference Notes

[1]The "Testwiseness Rating Scale" was designed for this study by Flippo, R., Borthwick, P., and Blanchard, J., 1981.

References

Callenbach, C. A. : The effects of instruction and practice in nonsubstantive test-taking techniques upon the standardized reading test scores of selected second grade students. Unpublished doctoral dissertation, Pennsylvania State University, 1971.

Crehan, K. D. Developmental aspects of test-wiseness. Paper presented at the annual meeting of the American Educational Research Association, New York, NY, 1977.

Diamond, J. J., Ayrer, J., Fishman, R. & Green, P. Are inner city children test-wise? Journal of Educational Measurement, Spring 1976, 14, 39-45.

Diamond, J. J. & Evans, W. An investigation of the cognitive correlates of test-wiseness. Journal of Educational Measurement, Summer 1972, 9, 145-150.

Dunn, T. F. & Goldstein, L. G. Test difficulty, validity, and reliability as functions of selected multiple-choice item construction principles. Educational and Psychological Measurement, 1959, 19, 171-179.

Flippo, R. F. Objective testwiseness module for education students. From Testwiseness. Rehoboth, MA: Twin Oaks, under contract.

Gibb, B. G. Test-wiseness as secondary cue response. Unpublished doctoral dissertation, Stanford University, 1964.

Jongsma, E. A. & Warshauer, E. The effects of instruction in test-taking skills upon student performance on standardized achievement tests. Final Report. New Orleans: University of New Orleans, Department of Elementary and Secondary Education, 1975.

Millman, J., Bishop, C. & Ebel, R. An analysis of test-wiseness. Educational and Psychological Measurement, 1965, 25, 707-726.

Oakland, T. The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. Journal of School Psychology, 1972, 10, 355-360.

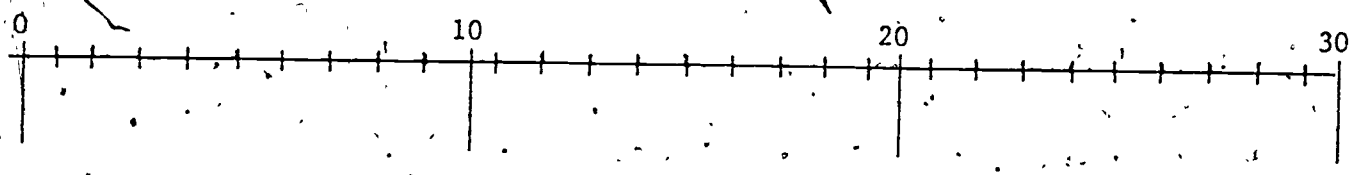
Ryan, T. A., Joiner, B. L., & Ryan, B. F. Minitab student handbook. North Scituate, MA: Doxbury Press, 1976.

Wahlstrom, M. & Boersma, F. The influence of test-wiseness upon achievement. Educational and Psychological Measurement, 1968, 24, 413-416.

Yearby, M. E. The effect of instruction in test-taking skills on the standardized reading test scores of white and black third grade children of high and low socioeconomic status. Unpublished doctoral dissertation, Indiana University, 1975.

TESTWISENESS RATING SCALE

by Flippo, Borthwick and Blanchard



Child doesn't seem to be able to get the answers correct. even when he/she seems to know the material all of the time.

Child seems to be able to get the answers correct some of the time.

Child seems to be able to get the answers correct most of the time.

Child seems to be able to get the answers correct even when he/she doesn't know the material all of the time.

Figure 1.

Table 1

Testwiseness Rating Scale:

Directions for Classroom Teachers

by Flippo and Borthwick .

1. Make a copy of your class roll.

2. Rate each child on your roll to show where you believe that child is in regard to his/her present testwiseness knowledge. You must select a number between 0 and 30. Please do not select the numbers 0, 10, 20 or 30.

3. Indicate the number selected next to each child's name on the roll.

4. After the children in your class have been randomly assigned to control (those who will not receive testwiseness training) and experimental (those who will receive testwiseness training) groups, indicate next to the number on the roll an E for experimental and a C for control.

5. If the class is ability grouped, please indicate low, middle, or high ability group also next to each child's name. If the entire class is either a low, middle, or high ability group, indicate that at the top of the roll. If the entire class is a mixed ability group, please indicate that at the top of the roll.

Table 2
Pre-Treatment Ratings on Testwiseness

Group Comparison	Median		p level
	Experimental	Control	
Second Grade-High Ability Transportation Unit	26.00	28.00	.0814 (NS)
Second Grade-Average Ability Transportation Unit	24.50	23.00	.6985 (NS)
Third Grade-High Ability Mississippi Unit	21.50	21.00	.8918 (NS)
Third Grade-High Ability Differences Are OK Unit	23.00	25.00	.2934 (NS)
Third Grade-High Ability US Postal Service Unit	21.00	21.00	.6261 (NS)
Fourth Grade-High Ability Regions of the US Unit	22.00	21.00	.2727 (NS)
Fourth Grade-Average Ability Japan Unit	27.00	27.00	.9611 (NS)
Fifth Grade-High Ability Tooth Care Unit	23.50	21.00	.5669 (NS)
Sixth Grade-High Ability Mexico Unit	21.50	24.00	.0793 (NS)

Table 3
Comparison of Unit Posttest Scores by Group

Group Comparison	Median		p level
	Experimental	Control	
Second Grade-High Ability Transportation Unit	100.00	100.00	1.000 (NS)
Second Grade-Average Ability Transportation Unit	97.50	92.50	.6985 (NS)
Third Grade-High Ability Mississippi Unit	94.00	82.00	.4025 (NS)
Third Grade-High Ability Differences Are OK Unit	82.00	79.00	.8182 (NS)
Third Grade-High Ability US Postal Service Unit	94.00	93.00	.8983 (NS)
Fourth Grade-High Ability Regions of the US Unit	99.00	98.50	.8625 (NS)
Fourth Grade-Average Ability Japan Unit	100.00	100.00	1.000 (NS)
Fifth Grade-High Ability Tooth Care Unit	89.00	85.00	.1921 (NS)
Sixth Grade-High Ability Mexico Unit	85.50	88.50	.5934 (NS)