Test scores were compared for two types of objective achievement tests—multiple choice tests and the recently developed Multi-Digit Test (MDT) procedure. MDT is an approximation of the fill-in-the-blank technique. Students select their answers from long lists of alphabetized terms, with each answer corresponding to a number from 001 to 999. The answer sheets are scoreable by computer. The results of previous research, as well as recent research in college undergraduate and high school classes, are both inconclusive. Some of the studies indicate advantages for the fill-in-the-blank style of test items, while other studies show no difference. No study has demonstrated that multiple choice testing resulted in superior test achievement. When students were questioned about their opinion of MDT, about half found it less desirable than multiple choice tests, yet preferable to open-ended fill-in-the-blank questions. About half considered MDT to be equal to or preferable to multiple choice tests. There was also some indication that study habits were different when students expected to be given MDT rather than multiple choice tests or traditional fill-in-the-blank tests. (GDC)
COMPARISON OF COGNITIVE ACHIEVEMENT IN OBJECTIVE TESTING:  
Multi-Digit and Multiple-Choice Tests

Dr. Paul S. Anderson  
Departments of Geography-Geology  
Illinois State University  
Normal, IL 61761  
(309) 438-7360

Dr. Eileen M. Kanzler  
Social Studies Department  
University High School  
Illinois State University  
Normal, IL 61761

(A paper presented to the conference of the National Educational Research Association (AERA), in Chicago, Illinois, on April 14, 1985.)

ABSTRACT

Two objective test methods are examined to determine which yields higher cognitive achievement. The Multi-Digit Test (MDT) procedure, available for only two years, is a computer scored approximation of fill-in-the-blank questions, which facilitates question writing and improves feedback to students and teachers. The students select their answers from long lists of alphabetized terms, each with a number up to 999 on the 3-digit format. Previous comparative research plus recent work specifically on the Multi-Digit Test are all similarly inconclusive. Some results show advantages for fill-in-the-blank style questions; other results show no difference; no study says that multiple-choice is superior for cognitive achievement. Recent development and availability of the Multi-Digit Test will permit others to research, and utilize this additional evaluation method.

I. Introduction and Background

This paper examines the effectiveness and influence of the Multi-Digit Test on cognitive achievement in comparison with the traditional multiple-choice test. The research described here focuses only on examination material which could be handled by either of the methods and also by the fill-in-the-blank method, i.e., when the answer is a discrete term or name. The objective is to be constructive, adding something new for specific applications while not detracting from the existing test methods, each with its own particular rationale.

Objective tests such as multiple-choice, true-false, and matching are frequently criticized as being too trivial or for encouraging guessing in a context of focus recognition. They are also faulted for taking too much of the teacher's time to compose properly. However, they are credited with the advantage of being easier and quicker to grade than the open response fill-in-the-blank tests which require recall. Whatever the advantages and disadvantages of objective tests, they are a mainstay of high school and college testing procedures and are the principal devices used in standardized tests. Yet clearly part of their content could be handled by the fill-in-the-blank method.

II. The Multi-Digit Test (MDT)

To diversify classroom objective tests, an instrument has been needed to evaluate large numbers of students while avoiding the guessing and recognition possibilities of the 5-10 foil multiple-choice test. At the same time the method should not require the time consuming manual grading of the fill-in-the-blank format. A new testing method which incorporates these features has been developed and in use since January 1983. It is called the Multi-Digit Test (MDT), wherein after reading the questions the students must have their answers in mind before referring to a list of several hundred alternatives arranged in alphabetical order. (See a sample list in Figure 1.) The code numbers of the selected answers are recorded
on an optically scanable answer sheet. A three-digit sheet presently in use can accommodate up to 999 code numbers. (See Figure 2.)

Students preparing for the MDT questions must study for recall rather than recognition. There are also other advantages. Once the answer lists are prepared the teacher has more flexibility and opportunity to compose and utilize a greater variety of questions. Computer scoring offers rapid grading and a full contingent of statistical analysis. Furthermore, the feedback to the students is greatly enhanced because the computer has a lexicon of responses that match the thousand alternative answers. A further discussion of the characteristics of the Multi-Digit Test was written by Anderson (1984a).

The Multi-Digit Test has been used with over six hundred students during five semesters, the majority being enrolled in a freshman-level World Geography course. As expected, student reactions are mixed. Unfamiliarity with the method, an additional sheet of paper on the desk, and the increased question difficulty (in contrast to 5-choice questions on the same subject matter) lead about half of the students to consider the

Figure 2: MDT answer sheet (reduced size)

Figure 2b: Reverse side of MDT form
MDT method to be less desirable than the traditional multiple-choice method. However, those students "...really prefer MDT when given the choice between free and full free-response questions (fill-in-the-blank).

About half of the students consider the MDT method to be equal to or preferred over the multiple-choice style. A sample of their comments include:

"...studied a great deal...shows that the individual knows the material or not, he can't guess..." S.B.

"...really have to know the information..." S.N.

"First time taking an Exam like this!!! I like it." G.T.

"I liked the MDT test system." T.H.

"...great idea" L.A.

"...liked the idea...more effective in determining if the students really know the information, you either know it or you don't." R.B.

"I think the MDT test method is very effective in that it contains the best aspects of both the multiple choice test and the fill-in-the-blank test method. It negates the problem of lucky guessing common to multiple choice test but at the same time allows recall of answers "on the tip of your tongue" by being able to see the name on a list. It also doesn't bring with it the problem of learning how to spell strange names correctly, a problem common to fill-in-the-blank tests." P.D.

"...good test method..." K.B.

"I like the MDT method. I feel it provides a much more accurate score when an evaluation of a student's knowledge is being attempted...If I know I would be faced with a MDT test...the motivation to study would be greatly enhanced. MDT tests will raise retention and lower unearned grades." R.E.

III. Research on the Comparison of Cognitive Achievement

A. Previous Research: Student comments and apriori reasoning suggest that the Multi-Digit Test should result in some advantages in cognitive achievement when compared with the 5-choice method. Previous research efforts in this area by Duchastel (1981) and Duchastel and Nungester (1982) indicate the efficacy of recall and consolidation-retrieval of the short answer (fill-in-the-blank) test over the recognition/discrimination pattern of the multiple-choice test. The same "mental review" thought process required in the short answer test is present in the MDT method. Other comparable studies (Sax and Collett 1968; Gay 1980) yield inconclusive or unrelated results. Ward (1982) reported that the response format made no difference in the pattern of correlations among a set of verbal aptitude tests. Such a variety of results suggest that further research is necessary.

B. Research with the Multi-Digit Test: Research by Anderson, Hill, Naim and Walters (1985) compared three test methods (including MDT) on the place name portion of a World Geography course. Their exploratory research used a Latin Square/recurring measures methodology with 188 students during a semester-long comparison. The students were always advised as to the test method used for each segment of their course. An announced post-test provided the final data. Retention of cognitive learning was highest when the students had studied for their turn with the fill-in-the-blank method, second with MDT, and lowest with multiple-choice, exactly as hypothesized. Although only the fill-in-the-blank method yielded a statistically significant difference (retention approximately eight to ten percent greater than in the other two methods), the results are considered to be conservative because of limitations in the research methodology.

Subsequent and on-going research has been undertaken by the co-authors (Anderson and Kanzler) during the Fall 1984 and Spring 1985 semesters within the setting of the laboratory high school at Illinois State University. The advantages of that environment have been discussed in an earlier paper by Anderson and Kanzler (1985).

Three of the five U.S. History classes in the laboratory school were taught by Dr. Kanzler and were utilized in the research. All students had previously pre-registered into classes of about twenty-five students in a random fashion based on availability, requirements, and electives. The pre-test and first-two tests of the semester were identical for all three
groups in order to ascertain any noteworthy differences between the classes. The means and other descriptive statistics were essentially the same, and the 11 AM class was selected as the experimental (MDT) group.

For each of the bi-weekly tests the MDT format questions posed to the experimental group were converted to multiple-choice 4-foil format for the two control groups. Five to ten multiple-choice questions where reasoning or judgment was a factor in the answer were also included on most tests. For the third test (first one with MDT), no advance notice on the change of format was given to the experimental group. As anticipated the experimental group achieved a much lower mean score (52%) than either of the two control groups (74% and 71%). That is a decline of 28% from the average correct scores of the control groups. It was readily apparent to the students that the MDT based on recall was a much harder, more demanding test. When the same test was subsequently administered in a multiple-choice 4-foil format to the experimental group, those students achieved the same (72%) result as the control groups. A few students complained about the difficulty and fairness of the MDT test, not realizing that difficulty is at the discretion of the instructor, who could also have given fill-in-the-blank tests.

Attempts to treat the MDT as just another but longer form of matching test had to be dispelled by enforcing a time limit only slightly longer than the average time needed to complete a multiple-choice test of equal length. The "search" for the correct answer by random scanning of lists or by elimination attempts is apparently a difficult habit to break.

After the first few administrations of the MDT tests, the students realized that they would have to modify their study habits to accommodate the recollection of factual evidence. Some resisted rhIs and dismissed the material as trivial, insignificant, and irrelevant. This too is at the discretion of the teacher and is not a criticism of the Multi-Digit Test. Since most students apparently rely heavily on the recognition factor present in multiple-choice tests, some object to the time and effort required by the demands of the recall strategy. Other students claimed that the change in study habits was impairing their performance on the multiple-choice section of tests where more judgment and reasoning were required. That was apparently not the case because an analysis of the multiple-choice segment of the final examination revealed that the students in the experimental group did as well if not slightly better than the control groups.

As the semester continued the experimental group made noticeable gains, at times attaining scores equal to those of the groups with the multiple-choice tests. However, their performance on the comprehensive final exam was disappointing (25% lower), but the sheer volume of information was probably a major factor. On one occasion the control groups were given both formats of the test with the MDT format coming first and being unannounced. Their MDT results averaged 42% lower than their own multiple-choice scores and 37% lower than the MDT scores of the experimental group. During the present spring semester, most of the students are still enrolled in Dr. Karizler’s second semester history classes. Further data collection, including a post-test in May, will yield additional results.

C. Summary: In general, the research efforts to date have not shown dramatic improvements in learning and retention that could be attributed to the Multi-Digit Test. On the other hand, no evidence even hints at any detrimental effect when MDT or open response methods are applied to appropriate subject matter. There are sufficient indications of possible benefits to merit further research. Very clearly, with questions which can be answered with a discrete term or name like those on the lists, the MDT method is more rigorous and students respond in a desirable way to that rigour. Although many students indicate that their study habits are modified, the qualitative and quantitative aspects of that modification have not been clearly identified.

IV. Conclusions and Future Work

The work on the Multi-Digit Test over the past two years indicates that it has definite applications in an important niche in computer-scored testing. Further variations of those applications will be found when additional disciplines give consideration to the MDT method. At present, MDT has only been used in selected courses of geography, history, home economics, and music. Virtually every discipline that has a
component of key vocabulary or content (like chemical elements) could utilize the Multi-Digit Test. Actual numeric answers of up to three digits (with the decimal point indicated in the stem of the question) provide new avenues for applications in a variety of quantitative subjects. Much more research on both the applications and the pedagogical issues is needed.

Work on the Multi-Digit Test is continuing at Illinois State University where larger samples and more disciplines are becoming involved. Other colleges and universities are invited to enter this research. An array of answer sheets, computer programs, long lists, and written guidelines are available for people interested in research on the method or interested simply in introducing it into their testing structures. The Department of Geography-Geology at Illinois State University is at present preparing for the dissemination of the materials so interested schools which commit to undertake research and evaluation of the Multi-Digit Test. A system of full or partial grants has been established to assist those schools or individuals in implementing the Multi-Digit Test. (Those interested should contact Dr. Anderson at the address given above.) It is anticipated that at some future conferences further presentations and possibly whole sessions could be devoted to the applications of the Multi-Digit Test.

Bibliography

Anderson, Paul S. (1984a), "An Introduction to the Multi-Digit Test (MDT)", Discussion Papers in Geography, No. 2, "Objective Testing in Geography", Old Dominion University, Norfolk, Virginia, pp. IV-1 to IV-18. (Originally presented as "Introduction to the Large-List Multi-Digit Test Procedure: With Examples From A World Geography Course", a paper at the West Lakes Regional Conference of the Association of American Geographers, Iowa City, Iowa, November 3-5, 1983.)

Anderson, Paul S. (1984b), "Applications of the Multi-Digit Test (MDT) Procedure for Teaching the Geography of Latin America", a paper presented at the Conference of Latin Americanist Geographers (CLAG), Ottawa, Canada, September 27-29, 1984. (A summary of this paper was published in CLAG Communication, Newsletter No. 50, December 1984, pp. 2-3)


Anderson, Paul S. and Eileen M. Kanzler (1985), "Laboratory Schools As A Unique Setting For Research: The Experimentation with the Multi-Digit Test (MDT) at Illinois State University High School", a paper presented at the conference of the National Association of Laboratory Schools, in Denver, Colorado, on 26-28 February 1985.


