DOCUMENT RESUME

ED 144 602 IR 005 230

AUTHOR Wood, Lewis Jay
TITLE Computer Assisted Test Construction in the BYU Library School.
PUB DATE Aug 76
NOTE 38p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Computer Assisted Instruction; Item Banks; Library Education; Post Testing; Pretesting; *Test Construction; *Testing; Tests
IDENTIFIERS Brigham Young University UT

ABSTRACT

Computer assisted test construction (CATC) is a new testing technique that seems to provide ease and flexibility for faculty members and students. The purpose of this paper was to verify that student test scores are not adversely affected by implementation of CATC. Two sections of a basic course in cataloging were tested for one semester. One section of 30 students was used as a control group while the other 20 students served as the experimental group. The students were pre-tested and then received class instruction. During the course of the semester, unit exams were administered. Exams for the control group were prepared by the course instructor. The experimental group's exams were compiled of random questions selected by the computer from the item bank. The analysis of scores from a post-test confirmed the hypothesis there would be no detrimental effect upon test scores for students tested with the CATC system. The pre-test used in the study is included in the appendix. (Author/JAB)

COMPUTER ASSISTED TEST CONSTRUCTION IN THE BYU LIBRARY SCHOOL

by

LEWIS JAY (BUD) WOOD

A Research Paper
Submitted to The
School of Library and Information Sciences
Brigham Young University
Provo, Utah

In Partial Fulfillment
of the Requirements of
L.I.S. 697

August, 1976

ABSTRACT

The problem treated in this paper was one of determination of effect upon student test scores caused by implementation of a computer assisted test construction (CATC) system.

Two sections of L.I.S. 528, a basic course in cataloging, were tested for one semester. One section was used as a control group. The other served for experimental purposes. The students were pre-tested and then each received class instruction. During the course of the semester, unit exams were administered to evaluate individual progress. At the conclusion of the semester, the students were post-tested.

It was found that students in both sections attempted to memorize the unit exams. Students in the control section were successful; while those in the experimental section were unable to do so.

The post-test analysis confirmed the hypothesis that there would be no detremental effect upon test scores for students tested with the CATC system.

PREFACE

This paper is written, not only with the intent of fulfilling requirements for LIS 697, but also in hopes that it, or a variation of it, might be used by the Brigham Young University (BYU) Department of Instructional Evaluation and Testing to help promote the use of computer assisted test construction (CATC) at the University.

The author, who serves as manager of Testing Services, takes full responsibility for the statements and judgments made herein, as they are based on his experience in Testing Services and his efforts in support of CATC at BYU since 1971. Gratitude is felt for the help and assistance of Dr. Adrian VanMondfrans, director of the BYU Department of Instructional Evaluation and Testing who permitted me time off work to complete my degree; to Dr. Merle Lamson, of the BYU School of Library and Information Sciences, for use of his LIS 528 classes; to the students of LIS 528, Winter Semester, 1976, for their willing (or unwilling) help; and to my wife, Elaine, and sons, Michael, Mark, Nathan, Scott, and Wendell, for four years of enduring while Dad went back to school.

TABLE OF CONTENTS

	Page
PREFACE	ii
LIST OF TABLES	v
Chapter	
1. INTRODUCTION	1
STATEMENT OF THE PROBLEM	1
HYPOTHESIS	1
DEFINITIONS	1
Item	1
Seed	2
Item Bank	2
Foil	2
Generation	2
Test Generator	2
2. CATC SYSTEMS--A BACKGROUND	3
3. ITEM BANK DEVELOPMENT AND TEST DESIGN: METHODOLOGY	8
THE CONTROL GROUP	10
THE EXPERIMENTAL GROUP	11
ANALYSIS OF THE TWO GROUPS--METHODOLOGY	12
4. ANALYSIS OF L.I.S. 528 TEST DATA	15
PRE-TEST RESULTS	15
UNIT EXAMS--TEST RESULTS	16
Retakes of Unit Exams	17

Multiple Retakes of Unit Exams 18
Unit Exam Summary. 18
POST-TEST ANALYSIS AND PRE-POST COMPARISONS. . 19
5. CONCLUSIONS AND SUGGESTIONS. 21
PRE-TEST RESULTS 21
UNIT EXAM RESULTS 21
POST-TEST RESULTS. 22
SUMMARY AND SUGGESTIONS FOR FURTHER STUDY. . . 22
APPENDIX 24
SOURCES CONSULTED. 30

LIST OF TABLES

Table		Page
1.	Pre-test Results	15
2.	Unit Exam Results	16
3.	Percentage of Students Retaking Exams	17
4.	Multiple Retakes of Exams	18
5.	Pre-post-test Comparisons	20

Chapter 1

INTRODUCTION

Statement of the Problem

Computer assisted test construction (CATC), when coupled with an out-of-class testing program, is a relatively new testing technique that seems to provide ease and flexibility for the faculty member and the student. It is important, however, before this program becomes widely used, to verify the assumption that student test scores are not adversely affected by its implementation. Such verification is the problem treated by this paper.

Hypothesis

The hypothesis tested in the following pages is that test scores from students involved in an out-of-class CATC testing program at Brigham Young University (BYU) will not be significantly lower than scores from students tested with traditional "teacher-made" examinations.

Definitions

Certain terms, as explained below, have specific and unique meanings as they are used in this paper. Such terms are defined as follows:

Item. An item is a test question, either written by a faculty member or made up by a computer, to be included in an examination.

Seed. A seed is a question in skeletal form that provides a framework upon which a computer will build to make up a test question. For example, "________ ________ was the ________ President of the United States." would be a seed that could be completed by the insertion of the words "Abraham Lincoln" and "sixteenth" to make a true-false question.

Item bank. An item bank is a collection of test questions or seeds, in machine readable form, from which items are selected to assemble a student test.

Foil. A foil is an alternative answer to an item. Given the question, "Who was the sixteenth U.S. President?", and the alternatives "George Washington," "Abraham Lincoln," and "John Kennedy," these three alternatives would be foils.

Generation. The act of assembling items to create a student exam is called generation.

Test generator. A computer program that assembles test questions to create a test is a test generator.

Chapter 2

CATC SYSTEMS--A BACKGROUND

A fairly recent development at Brigham Young University is the development of a computer assisted test construction facility administered by the University's Department of Instructional Evaluation and Testing. BYU's CATC programs involve the computer in test construction and analysis as well as in grading and record keeping to assist the faculty member in his task of student evaluation.

According to Gerald Lippey, manager of Advanced Instructional Applications Development in the Data Processing Division of IBM, "The most noteworthy changes in educational testing during the past few decades have been those which resulted from technological progress."[1] The late 1940's and 1950's saw the advent of machines developed to quickly and efficiently score student examinations. Since that time, machine scoring technology has advanced so that several thousand exams can now be scored each hour by a single, relatively untrained, person. As the scoring is done, data can be gathered to provide somewhat sophisticated statistical measures of not only the exam, taken as a whole, but also of each alternative or foil on each question of the exam.

---

[1]Gerald Lippey, ed., Computer Assisted Test Construction (Englewood Cliffs, New Jersey: Educational Technology Publications, 1974), p. 3.

Computer assisted test construction came to the Brigham Young University campus in 1971, when Testing Services (officially, the Department of Instructional Evaluation and Testing) developed an initial test generator program to select items randomly from an item bank of questions provided by the History department. In the past four years, CATC testing at BYU has grown from a yearly total of 2,000 exams administered, to a projected 1975-1976 academic year total of 175,000 examinations.[2]

Exactly what is CATC? Computer assisted test construction has many facets, yet there are several commonalities amongst all systems as they have developed in the United States. The systems include an item bank for examination purposes. In most cases, the bank itself is stored in the computer, although at times, only indices to items, or item seeds are machine readable.

Most of the items in existing CATC systems are objective. However, they need not be. BYU's test generator program can handle not only objective questions (of which true-false are simply a subset), but also matching, short answer, and essay questions. CATC involves, in every case, the use of a computer to select items from the item bank. The algorithm for selection is generally quite simple in that items are basically classified before selection, by

---

[2]Lewis J. Wood, "Request for Approval of Computer Hardware Procurement" (Provo, Utah: Brigham Young University Department of Instructional Evaluation and Testing, 1975), p. 4.

subject and/or other measures such as difficulty or grade level. The instructor then sets parameters to be used by the generator program, indicating test length, subject composition, difficulty, etc.

CATC, especially when linked to an out-of-class testing program such as that offered at BYU, seems to offer several advantages over a conventional testing program. Use of the computer to gather item response information and statistical measures on each foil allows for continual item improvement, for example.

Because item randomization virtually assures test security by preventing "leaks" of test forms, out-of-class testing becomes practical. Such testing procedures provide increased lecture time within the classroom. For the student, out-of-class testing provides flexibility to his schedule, allowing him to avoid "heavy" days wherein he would normally have several exams.

The proper development of item banks and their increasing usage provide an opportunity for sharing of items among several faculty members. While this practice is possible without a machine readable item bank, the existence of the bank makes it practical and easy to share questions of potential worth among faculty members.

Cost is another major advantage in a well-designed CATC system where multiple forms of several equivalent examinations can be quickly, easily, and inexpensively generated. "Computer costs per printed test last year [at BYU] totaled $0.4334. . . ."[3] Brigham Young University's CATC program has been designed so that fourteen to sixteen students use the same form of each exam. This sacrifices some test security, but also cuts costs. The per-student test generation cost to the department last year was only $0.0271.[4] In a recent study performed by the author for the Department of Instructional Evaluation and Testing, the cost of generating student examination, administration of the tests, scoring, and reporting back cumulative grades to the faculty was found to be $0.3818 per student per test. This compares favorably to the $1.72 per student per test cost found for conventional test preparation, administration, scoring, and recording.[5]

In order to take advantage, then, of the CATC program, what steps must be taken? The first step is the selection or development of a test generator program. Points to consider here include minimization of costs, criteria for item identification and selection, file organization, and item response feedback for question evaluation. Each of these questions has to be answered by the systems analyst who will design

---

[3] Lewis J. Wood, "Establishing a CATC System: Where to Begin" (paper read at the Second Annual Conference on Computer Assisted Test Construction, Atlanta, Georgia, October 13, 1975), p. 5.

[4] Ibid.

[5] Lewis J. Wood, "Preliminary Analysis of Costs and Revenues for Modular Testing" (Provo, Utah: Brigham Young University Department of Instructional Evaluation and Testing, 1975), p. 5.

the test generator. Inasmuch as the BYU program is already written, it now is sufficient to mention these points and then pass on to the particular problem at hand: the development of an item bank to be used in L.I.S. 528, the cataloging course in the BYU School of Library and Information Sciences.

Chapter 3

ITEM BANK DEVELOPMENT AND TEST DESIGN: METHODOLOGY

Before an adequate item bank could be developed for L.I.S. 528, an item analysis had to be run on the items previously used in the class's five unit examinations. This analysis provided not only frequency response information on each item foil, but also a point-biserial correlation for that foil. The item analysis program then used that correlation figure to perform a question evaluation which rated the item's discrimination record from "A" (a discriminating item) to "E" (an ambiguous one). Based upon this question evaluation, the item pool was culled of all "D" and "E" questions before it was put into the machine readable format.

Because of the design specifications of TESTGEN, the test generator program in use at BYU, the items for the pool were keypunched to eighty column cards. The first four columns of each card, as well as the last five columns of the last card in the question were used for control and identification purposes. The item could be of any length, however, as any number of cards could be used for the question text. As a backup, to be used in cases of extreme machine error, the entire item bank was also stored on tape as well as on disc. The disc file was the actual "generation file" from which the items were randomly selected.

The item pool for use in LIS 528 was strictly multiple choice in format. Not only is this the traditional approach to unit exams for this course, it is also the easiest type of question for which to quantify test results. Consequently, no change in item format was made. However, using a technique first reported by Denney in 1973,[1] but independently developed at BYU in 1971, the multiple choice questions had a unique difference. Several incorrect responses were loaded into the item bank for each question and the generator program randomly selected not only the item to be used on the examination, but also the responses to be printed with that item. Thus, while two forms of the exam could contain the same item, the foils to that item could be unique.

One problem of particular significance in the development of any machine-readable item bank is that of determining the size of the bank. In other words, how many questions are needed in an item bank? "Many theoretical and practical factors are involved in the final decision. The number of items must be adequate to cover the subject matter. . . ."[2] Lippey judged that about fifty items are needed per class hour of presented material. Donald Jensen, professor of psychology at the University of Nebraska, postulated that

---

[1] C. Denney, "There is More to a Test Pool than Data Collection," Educational Technology, 13 (1973), 19-20.

[2] Gerald Lippey, ed., Computer Assisted Test Construction (Englewood Cliffs, New Jersey: Educational Technology Publications, 1974), p. 48.

about ten times as many relevant items are needed in the bank as will appear on each test. It is interesting to note, however, that the two methods yield somewhat similar results. Jensen administered about eight examinations in a semester using a total of 2,400 items. Lippey's rule applied to a similar course meeting forty-five times yields 2,250 items.[3]

For the above reasons, it was originally planned to apply the CATC approach to the LIS 528 quizzes rather than to the unit exams. Inasmuch as this was not possible, the available items from the unit exams were used. These items totaled only approximately twice the number of items to be selected for each test generated.

Once the item pool was prepared and edited, it then became possible to set up testing programs for two separate sections of LIS 528, to perform pre-test analysis, and to run the students through the programs, using one section for control and one for experimental purposes. The following paragraphs explain how the program worked.

The Control Group

The students in the control group attended class in the traditional manner. Weekly quizzes on cataloging, written by the class instructor, were administered in the modular test center located in the Grant Building on the BYU campus. The center was open from 8:00 AM until 6:00 PM on Mondays; 8:00 AM until 8:00 PM, Tuesdays through Fridays; and 9:00 AM until 1:00 PM on Saturdays. Each of the five unit

---

[3] *Ibid.*, p. 49.

exams, also written by the class instructor, was, in a like manner, administered. Each quiz and unit exam had to be passed with a score of 81 percent or above. Failure to do so necessitated a student retaking the test. The retake quiz was made up of questions similar to those on the first quiz. The retake exam, as has been traditional, was made up of the same items, printed in a different order. To keep track of which version of an exam or quiz was taken, each student was issued a testing card by Testing Services. As soon as the student finished his test, it was hand-scored and he was then able to review the material on the test to see where mistakes were made.

The Experimental Group

The students in the experimental group also attended class in the traditional fashion. The same weekly quizzes administered to the control section were also administered to students in the experimental group. Methods of test administration were the same for both groups as were modular center hours.

The experimental section was also administered five unit exams, but these exams were composed of items randomly selected from the item pool. As in the control group, each student had to retake a quiz or exam if he failed to attain a minimum of 81 percent. Retakes of quizzes for the experimental class were conducted the same as for the control class. Retakes of the unit exams, however, were different inasmuch as the retake exam was another randomly selected (instead of simply re-ordered) test. Compared with the student's first version of the exam, some items were expected to be the same; others would be the same question with a different set of alternatives; while other items would be completely different. Item order was also randomized, in that the first question on form "A" was not necessarily the first question on form "B", if, in fact, it appeared on form "B" at all. Testing Services personnel recorded which forms of the quizzes and unit exams were taken by students in the experimental group by again using a testing card, as in the case of the control group. In order to separate the two groups and to insure the proper examination was given to each student, the testing cards were color-coded.

As soon as the student in the experimental section finished his paper, he took it to the control clerk where it was hand-scored. He was then able to look over his exam and determine where he made his mistakes.

Analysis of the Two Groups--Methodology

Assuming a normal distribution of students as they enrolled at the beginning of the semester, the students in the control and experimental sections of L.I.S. 528 were thought to be equally "unknowing." To verify this assumption, a pre-test, made up of questions selected from the unit exams and weekly quizzes, was written by the author of this paper and is included in the appendix. This pre-test was administered to students in both sections. A t-test was run against the scores to verify that no difference existed at the 0.01 level of significance.

At the conclusion of the semester, the same pre-test was again administered and these "post" test results were measured for statistical significance. Gains in both groups in pre- vs. post-testing were anticipated, but expections concerning differences of post-test scores between the two groups were not known. Higher post-test experimental scores or approximately equal post-test scores between the two groups should be indicative of a significant advantage of the CATC or experimental approach over the traditional testing due to ease of testing and grading for the faculty member. Significantly lower post-test experimental scores should, on the other hand, indicate the experimental approach had been detrimental to the students' learning processes.

Due to the required minimum passing score of 81 percent for both sections, t-tests were not run on the unit exam scores, but the hypothesis was made concerning these scores that the students in the control section would simply "memorize" the tests, while those in the experimental section would not be able to do so. Assuming this to be the case, the author anticipated that the highest mean scores for the experimental section would be lower than comparable scores for the control section, while the standard deviation (s.d.) of these scores for the experimental section would be greater. In other words, since the students in the experimental section were unable to memorize the tests, their scores would be lower and the curve would be more spread out than the curve of the control section.

If, as has been hypothesized, students have attempted to memorize the unit exams, it would also be expected that the greatest number of multiple retakes (third or fourth attempts at passing) would occur in the experimental section on Test 1, where the students anticipated returning to retake the same exam. As it became apparent that item memorization would not be helpful, increased study, previous to taking an exam, would become more appropriate and, consequently, it might be seen that total retakes for the experimental section would drop to less than that of the control section as the testing continued through the semester.

Chapter 4

ANALYSIS OF LIS 528 TEST DATA

Pre-test Results

Approximately fifty-five students enrolled in LIS 528 for Winter Semester, 1976, but 10 percent dropped the class before it actually began. Of the fifty students remaining who began instruction and actually took the pre-test, thirty were in Section 1 and were designated as the control group. The remaining twenty students enrolled in Section 90 and constituted the experimental group of students.

The t-test results of the pre-test can be seen in Table 1, below.

Table 1

Pre-test Results

Group	N	Mean	S.D.	Significance
Control	30	12.833	2.437	-
Experimental	20	12.300	2.003	-
Cross Test Significance				None

As can be seen, while the mean score of the control group exceeded that of the experimental section, the observed difference was not considered to be statistically significant.

This result verifies the assumption that the two sections started the course at the same point. In other words, the two groups were, in fact, equally "unknowing."

Unit Exams--Test Results

As can be seen from an examination of Table 2, analysis of the data collected from the unit exams is interesting.

Table 2

Unit Exam Results

Test	First Attempt		Highest Attempt		Average Retake Gain
	Mean	S.D.	Mean	S.D.	
Control Group					
1	75.88	11.97	90.49	6.78	28.19
2	77.44	13.24	93.15	5.76	28.27
3	80.22	9.66	90.82	6.36	23.86
4	83.65	8.31	88.46	7.31	25.00
5	85.63	9.22	90.25	4.78	23.13
Experimental Group					
1	70.25	12.30	88.25	7.95	28.50
2	73.75	8.35	84.75	5.80	22.00
3	77.25	12.30	84.75	6.00	18.75
4	82.00	12.90	89.50	8.00	15.00
5	84.25	8.55	85.75	7.65	25.00

In four of the five cases, the experimental students scored lower, after retaking the unit exams, than did the control students. Also, in four of the five cases, the s.d. of the control group was less than that of the experimental section. In three of the cases, the average gain over initial scores obtained by control students as they retook their exams exceeded the gain made by experimental students. In one case, this difference was a full ten points. In the two cases where this situation reversed itself, the experimental gain over the control gain was less than two points. (In fact, on Test 1, the experimental gain was less than one-third of one point better than the control gain.)

Retakes of Unit Exams. The percentage of students retaking exams was quite comparable for four of the five unit exams. Only on Test 5 did the experimental students drastically differ in the number of retakes. Table 3 shows this data.

Table 3

Percentage of Students Retaking Exams

Test	Control Group	Experimental Group
1	56	57
2	56	48
3	44	43
4	19	22
5	20	9

Multiple Retakes of Unit Exams. Multiple retakes of unit exams were significantly different for the two groups. Only 1.08 percent of the total exams taken by the control group were multiple retakes, while 5.37 percent of the experimental exams fell into this category. Fifty percent of all experimental multiple retakes occured on Test 1, but this number declined steadily until none of the students in the experimental section had to take the fifth test more than twice. These retakes, as a percent of the number of students initially taking the tests and of the actual numbers of multiple retakes recorded are shown in Table 4.

Table 4

Multiple Retakes of Exams

Test	Control Group (n=27)		Experimental Group (n=20)	
	% of Class	# of Retakes	% of Class	# of Retakes
1	3.7	1	20.0	4
2	0.0	0	10.0	2
3	0.0	0	5.0	1
4	0.0	0	5.0	1
5	3.7	1	0.0	0

Unit Exam Summary. In summary, the analysis of unit exams scores shows slightly higher retake means for the control section than for the experimental group. Conversely, experimental s.d.'s are smaller. Retake percentages tended to be quite close, until Test 5, where the experimental group as a whole retook the exam less than half the number of times that the control group did. Analysis of the multiple retakes of exams shows this happened almost five times as frequently (on a per capita basis) in the experimental class as in the control section. Exactly one-half the experimental multiple retakes took place on the first unit exam.

Post-test Analysis and Pre-post Comparisons

The sample sizes in the control and experimental groups declined during the semester to where only 41 students took the post-test. Seventeen of these were in the experimental section and the remaining 24 were in the control group.

As can be seen in Table 5, there was a considerable gain for each section over its pre-test scores. T-test analysis of this gain showed it to be significant at the .001 level of confidence. A similiar analysis of post-test means between the control and experimental sections showed there to be no significant difference between the two groups. A close examination of the mean gain for each section on the post-test over its pre-test score shows the students in the experimental section outscored their counterparts in the control group by 0.849 points. An analysis of covariance was run in an attempt to determine if this more sensitive test could find significance in this gain, but none existed at the .05 level.

18

Table 5

Pre-post-test Comparisons

Group	Pre Mean	Pre S.D.	Post Mean	Post S.D.	Mean Gain	Significance
Control	12.833	2.437	19.625	4.332	6.792	.001
Experimental	12.300	2.003	19.941	4.235	7.641	.001

A summary tabulation of the pre-post-test data concludes that the two groups started together and, while mean gains in post-test scores over scores on the pre-test were significant at the .001 level, the two groups had no significantly different learning experiences. This is in spite of the observed difference in mean gain scores where the experimental section did better.

Chapter 5

CONCLUSIONS AND SUGGESTIONS

Analysis of the data collected has proven to be quite conclusive as follows.

Pre-test Results

The t-test showed the two groups to be equally "unknowing" at the beginning of the semester. Thus, no adjustment of test scores had to be made in order to make post-test comparisons.

Unit Exam Results

As anticipated, retake scores from the control section were higher and more closely clustered than were the same scores from the experimental section. This, coupled with the retake patterns of the two groups and the excessive multiple retakes on Test 1 for the experimental section, provides conclusive evidence, in the author's opinion, of attempted test memorization on the part of both groups. In a like manner, the data suggest the experimental approach of random item selection to be fairly effective in combating this attempt, with resultant increased study on the part of those students in the experimental section.

Post-test Results

The experimental section showed more gain over pre-test scores than did the control section, but this gain was not significant at the .01 level. The fact that the experimental students did at least as well as the control students is significant, however, in that it verifies the primary hypothesis of this paper that scores of students involved with CATC testing should not be lower than the scores of students not so involved.

Summary and Suggestions for Further Study

The data gathered and presented in this paper presents a strong case, in the author's opinion, for CATC testing. The tests are easier for the instructor to produce, after the data bank is prepared, than are conventional exams. They also seem to have some advantages over the conventional exam if that exam is to be retaken to measure a student's learning and growth. Test memorization becomes impractical in such a CATC environment and this, in and of itself, can lead to increased study. Test security also is less of a problem, as students find no reason to pass questions to their peers.

More importantly, from the viewpoint of the class instructor, the data suggest there is no decrease in the student's learning due to CATC testing. On the contrary, it might just provide the stimulus for further study and a resultant enriched learning experience.

Thinking of this paper as a pilot to guide those to follow, expansion of the project so as to involve several hundred students over multiple semesters in an extensive evaluation of the impact of CATC testing might be fruitful. Preliminary analysis of this impact indicates the emergence of a new tool, which may prove to be quite useful in enriching the educational process for students involved in its use.

APPENDIX

GRADUATE SCHOOL OF LIBRARY & INFORMATION SCIENCES

DO NOT WRITE ON THIS TEST

LIS 528
Organization and Processing
of Materials

M. Lamson
Winter, 1976

PRE-ASSESSMENT INSTRUMENT

Please code your name and social security number in the appropriate boxes on the answer sheet.

GOOD LUCK!

1. What is the best way to catalog maps?

   a) use LC
   b) use AACR
   c) use American Geographical Society
   d) there is no one "best" way
   e) none of the above

2. Which publication seems to be the only one dealing with county and municipal items?

   a) monthly catalog
   b) checklist of State Documents
   c) PAIS
   d) Municipal yearbook
   e) none of the above

3. What is the major problem in cataloging music with generic titles?

   a) establishing the standard title
   b) no particular major problem
   c) establishing the composer
   d) none of the above--there is a problem, but it's not here

4. What is the difference between a full score and a miniature score?

   a) none
   b) size
   c) use
   d) none of these--there is a difference, but it's not here

5. The use of color-coded cards is recommended for use in cataloging media.

   a) true
   b) false

6. What is subject cataloging?

a) Subject cataloging deals with making the entire catalog card
b) Subject cataloging deals only with establishing the main entry
c) Subject cataloging deals only with deciding upon which subject headings to use

7. What are the two major categories of biography?

a) collective - individual
b) lives of persons - as a form of writing
c) collective, individual - ad hoc
d) as a form of writing - ad hoc
e) none of the above

8. The field of literature has two classes of materials which must be distinguished carefully. They are

a) belles-lettres - collections
b) collections - works about literature
c) belles-lettres - individual literature
d) work about literature - belles lettres
e) none of these

9. There are several uses for a shelf list. Which of the following is not one of them?

a) protection against duplication of a call number
b) buying guide
c) inventory control
d) record of achievement
e) aid in classification

10. What constitutes a "set of cards?"

a) main entry card, plus one card for each tracing
b) main entry card, plus one card for each tracing plus shelf list
c) main entry card and shelf list
d) none of the above

11. The LC Subject Heading List, beyond simply being a subject list, can also serve as

a) quasi-relative index to LC
b) a finding device
c) a name file
d) none of these

12. What edition is unabridged DC now in?

a) 14th
b) 17th
c) ~~18th~~
d) 20th
e) none of the above

13. LC traces itself back to which great philosopher?

a) Aristotle
b) Plato
c) Spencer
d) Cutter
e) none of the above

14. You have been using LC Class N: Fine Arts as a text; Which of the following subjects are included in DC 700's, but not in LC Class N?

a) Music
b) Photography
c) Graphic arts
d) a, b
e) a, b, c

15. What seems to be a major difficulty with use of almost any classification scheme?

a) overlapping of subject areas
b) no real major difficulty
c) language problems
d) none of the above

16. In the tracings on a card, subject entries:

a) precede other added entries
b) go behind Roman numerals
c) follow the other added entries
d) (a) and (b)
e) none of the above

17. What is the purpose of a <u>See also</u> reference?

a) To direct a reader from a non-used heading to a used heading
b) To provide historical kinds of information for the user of the card catalog
c) Both (a) and (b)
d) To direct a user to material related to the heading consulted
e) None of the above

18. What is the purpose of a Uniform Title?

a) To make added work for a cataloger
b) To bring together all catalog entries for a given work for which various editions, translations, etc. have various title.
c) To provide a method for standardizing title entries
d) None of the above

19. What is the entry work for the Holy Bible?

a) The Bible
b) The Holy Bible
c) Under name of translator
d) Bible
e) None of the above

20. Generally speaking, when two corporate bodies have the same name, but are located in different places, then:

a) some arbitrary device is used to distinguish between them
b) the name of the place is added
c) one is entered under place; the other is entered under its corporate name
d) hopefully such will not happen
e) none of the above

21. What is the primary importance of MARC?

a) large data base
b) communications device
c) networking device
d) standardization
e) none of the above

22. How are MSS housed?

a) In boxes
b) on the regular shelves
c) In acid-free manila folders
d) none of these

23. What is the major difficulty in establishing Chinese personal names?

a) Language impossible
b) one person may have several names which are all legitimate to use
c) none
d) few records available
e) none of the above--there is one, but it isn't listed above

24. What are holographic manuscripts?

a) typewritten
b) printed
c) handwritten
d) dittoed
e) none of these

25. Day-books, journals, diaries would be categorized as:

a) printers' copy
b) MSS written before the invention of printing
c) author's first drafts
d) correspondence not written for publication
e) private papers

SOURCES CONSULTED

SOURCES CONSULTED

1. Books

Lippey, Gerald, ed. Computer Assisted Test Construction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1974.

2. Periodicals

Denney, C. "There is More to a Test Pool than Data Collection," Educational Technology, 13 (1973), 19-20.

Toggenburger, Frank. "Classroom Teacher Support System," Educational Technology, 13 (1973), 42-43.

Epstein, M. G. "Computer Assisted Assembly of Tests at Educational Testing Service," Educational Technology, 13 (1973), 23-24.

3. Conference Proceedings and Papers

Baker, Frank B. "A Conversational Item Banking and Test Construction System," Proceedings of the 1972 Fall Joint Computer Conference, 41 (1972), 661-667.

Prosser, Frank, and Donald D. Jensen. "Computer Generated Repeatable Tests," Proceedings of the 1971 Spring Joint Computer Conference, 38 (1971), 295-301.

Wood, Lewis J. "Establishing a CATC System: Where to Begin." Paper read at the Second Annual Conference on Computer Assisted Test Construction, Atlanta, Georgia, October 13, 1975. (Proceedings yet to be published.)

4. Unpublished Materials

Wood, Lewis J. "Preliminary Analysis of Costs and Revenues for Modular Testing." Provo, Utah: Brigham Young University Department of Instructional Evaluation and Testing, 1975.

________. "Request for Approval of Computer Hardware Procurement." Provo, Utah: Brigham Young University Department of Instructional Evaluation and Testing, 1975.