A STUDY OF PEDAGOGICAL APPROACHES TO TEACHING PROBLEM SOLVING

by

BRIAN LYN SNYDER

B. A., Bethany College, 1977

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1985

Approved by:

Major Professor

TABLE OF CONTENTS

THIS BOOK CONTAINS NUMEROUS PAGES WITH ILLEGIBLE PAGE NUMBERS THAT ARE CUT OFF OR MISSING.

THIS IS AS RECEIVED FROM THE CUSTOMER.

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

The author wishes to express appreciation and gratitude to many kind and encouraging friends. To:

Dr. Elizabeth Unger, my advisor and friend. For enduring kindness, patience, understanding, and professional guidance during the preparation of this paper.

Dr. Raja Nassar, who served as my statistical consultant and was a ready source of information and help.

Dr. Richard McBride, for being there when questions needed answers and for being a friend.

Airika and Hannah, my daughters, whose ever-ready "help" in organizing, writing, and typing was joyfully shared with me.

My family and parents, whose faith and love was limitless...

Barbara, my wife, whose love, induring patience, and encouragement contributed to the success of this paper.

Chapter 1 - Introduction

The recent controversy surrounding the quality of education in America is anything but new. The issues of what should be taught and by what methods has been a subject of discussion since the days before the American Revolution. Even though education has been a topic of discussion for many years and the "ideal" solutions have been many, the one consistency has been the effort of all involved to make the American educational system the best that it can possibly be. One approach to attaining this goal has been the development of better instructional methods.

The search for the ideal instructional methodology has been a never-ending process for colleges and universities. In the recently released report, "A Nation at Risk," the report states that, "The necessity of determining the definitive methodology is imperative in order to improve upon our educational system."[1] It is with this in mind that a brief description of the differences among the faculty surrounding the effectiveness of the two methodologies used in teaching Algorithmic Processes (CMPSC300) offered at Kansas State University. During the 1982-83 and 1983-84 academic years, two distinct methods of teaching this course were used. The goals of the course include the development of the student's problem solving skills, as related to computer science problems, and algorithm development. The reason for the concern on the part of the faculty is that the course, Algorithmic Processes, is one of the cornerstones of the Computer Science degree. If a solid foundation is not obtained in Algorithmic Processes, a student is in a poor position to successfully complete the remaining required computer science courses.

1. "A Nation at Risk: The Imperative for Educational Reform," The National Commission on Excellence in Education, U.S. Department of Education, April 1983.

Normally, the methods that an instructor employs in the classroom is left to the discretion of that individual. Because of the importance of this course to the entire program, the faculty has a vested interest and a concern for high quality education in this course. As with any issue where there is a choice, positions has been taken supporting each methodology by faculty members. Aside from formal and informal discussion, no study has been made of the issue. It is, therefore, the purpose of this paper to explore the "problem" from a statistical point of view in an effort to determine if there is an advantage to be gained by using one of the methods over the other and to offer suggestions for improvement.

The first method, hereafter referred to as "Method A", can be categorized as the "Discovery Method" or in more precise terms "Essentialism".[2] In this methodology, the instructor has a definite goal or concept in mind that is to be reached during one or more class sessions. The key is that the students are usually not informed of this goal. Then, through dialogue with the instructor, the class is lead to a "discovery" of the "new" concept. This methodology requires that the instructor carefully control the discussion so that the desired goal is achieved.

The second methodology, "Method B", has as its focus, the acquisition of knowledge through an analysis of data. More commonly referred to as "Information Processing,"[3] in this methodology the students are viewed as an active participant versus a passive recipient. In this strategy, students are presented with a new idea and then through the use of various testing techniques, comparisons with known concepts, the student gains an

2. Teaching with Charisma, Lloyd Duck, Allyn and Bacon, 1981.
3. Strategies for Teachers -- Information Processing Models in the Classroom., Paul D. Eggen, Donald P. Keuchek, and Robert Harder, Prentice-Hall, Inc., 1979.

indepth understanding of the concept.

The general approach to this problem involves a) ascertaining the skill development of the student prior to taking Algorithmic Processes and b) gathering performance information on the student in Algorithmic Processes and succeeding Computer Science courses. The variables used to measure skill level prior to taking Algorithmic Processes were determined through discussions with various faculty and other University personnel. With the help of a statistical consultant, various statistical tests were run on this data to discover if either method is superior to the other or if both methods have approximately the same effect.

This report has been divided into four chapters. The process of gathering the data and the experiments used to analyze the data are described in chapter two. Chapter three details the results of the statistical testing. Chapter four discusses the implications of the results and outlines some areas of possible future study.

Chapter 2 - Experiment

The initial set of student observations, identified by student identification number, was gathered from the list of all students who had taken Algorithmic Processes within the 1982-83 and 1983-84 academic school years. Applying the restrictions listed below, in Table 2.1, and by using a list of Computer Science majors this original set of observations was reduced to the working set.

1) Computer Science undergraduates - The concern of the faculty is for the undergraduate majors and not so much for majors outside the department or graduate students.

2) Completion of one or more of the three follow-up courses (CMPSC305, CMPSC405, CMPSC460). Without completion of any of these courses, the application of the statistical tests would not be apropos. The study involves performance in succeeding Computer Science classes.

Table 2.1 Restriction Criteria for Study Subjects

Meetings with faculty members and the statistical consultant resulted in a set of independent variables which might have had an effect upon the performance of a student in the Algorithmic Processes course. For the most part, these variables were ones that were believed to represent the development of a student's logical and reasoning facilities. A complete listing of these variables is given in Table 2.2.

1) Sex of the student.

2) Prior course work in Biochemistry, Chemistry, Mathematics, Computer Science, Physics, Electrical Engineering, and Philosophy.

Table 2.2 Independent Variables Related to Student Performance in Algorithmic Processes

The process of collecting the data involved the cooperation of several offices and departments on the campus of Kansas State University. The first step was to obtain permission to locate and gather the necessary data. This was accomplished by receiving an opinion from the Chairman of the University Committee on Research Involving Human Subjects that the data was pre-existing and that the Department of Computer Science could collect the data. The one restriction placed upon such a collection was that the data could in no way identify a given student. The next step was to receive permission from the University Registrar to access their database of student records. Due to the necessity for security of student's records, the Registrar was reluctant to allow access to the Student Information System. The solution reached was to grant permission for the the data to be gathered from the Center for Student Development, which has tape copies of student records. Table 2.3 illustrates the record format of the final form of the data that was used during the statistical testing.

entry #	description
1	sex of the student
2-54	independent variables
55	if method A then 0 else 1
56	if method A then 1 else 0
57	if method B then 1 else 0
58	semester of CMPSC300 (coded)
59	score for CMPSC300
60	if method A then score CMPSC300
61	if method B then score CMPSC300
62	semester of CMPSC305
63	score for CMPSC305
64	semester of CMPSC405
65	score for CMPSC405
66	semester of CMPSC460
67	score for CMPSC460

Table 2.3 Format of Records Used in Study

A complete listing of the data is given in Appendix A. The dates for all

of the entries have been removed so that it would not be possible to identify an individual student.

The first test performed was a comparison of means for the Algorithmic Processes scores for the two methods. This was necessary in order to determine how these scores would be represented in the next test. A significant difference in the means would signal the need to exclude these scores in some of the later tests. This exclusion was due to the fact that Algorithmic Processes had been taught by two instructors, each using their own grading criteria. The results, shown in Figure 2.1, clearly demonstrates a difference of means too great to ignore. Consequently, instead of using the scores earned by the students in Algorithmic Processes, the values one (1) and zero (0) were used. An entry of one (1) signifies that the student had been exposed to that particular method, while a zero (0) indicates they had not been. This alleviated the problem of unduly influencing the remaining statistical tests.

Variable	N	Mean
S300	191	83.937
AS	98	80.633
BS	93	87.419

Figure 2.1 Mean Scores

The first test involved checking to see if either course had an effect on any of the selected follow-up courses. Regression Analysis was the selected tool using the SAS[1] Stepwise procedure. The model equation used

1. SAS Institute Inc., Cary, North Carolina (release 82.3 at Kansas State University).

in the analysis of the observations is listed below and was the same for each of the follow-up courses.

<course> = sex <all the X variables> A B S300

The variables A and B are defined per entries 56 and 57, respectively, of Table 2.3. The results of this, and the remaining tests, are presented in chapter three.

The choice for the next test was the SAS GLM procedure which would show whether the effectiveness of the two methods, indicated by the intercept of the regression equations, was significant or not. The remaining test were performed on a split data set where one set had the observations from Method A and the other those from Method B. This division of the data set allowed the removal of the variable A from the regression equations. The reason for using the regression equations selected by the Stepwise procedure is based, in part, on the findings of Draper and Smith, "We believe this to be the best of the variable selection procedures discussed and recommend its use."[2] A result of the GLM procedure would be an indication of the probability of the intercept being significantly different than zero.

The last step was to select the proper plan to compare the effectiveness of methods A and B of the selected follow-up courses. The following hypothesis was put forward to be used on each course.

---

2. Norman Draper and Harry Smith, Applied Regression Analysis, John Wiley & Sons, Inc., 1966.

$$\text{null hypothesis } H_o: T_a = T_b$$

$$\text{alternative hypothesis } H_a: T_a \neq T_b$$

The values T_a and T_b measure the effects of each method on each of the selected courses. The statistical test used to check the null hypothesis was the Student's t Test where

$$t = \frac{T_a - T_b}{\text{standard error of the differences}} \quad \text{with } n_a + n_b - k - 2 \text{ d.f.}$$

The method of computing the standard error of the differences, as outlined in Figure 2.2, is cited by William G. Cochran[3] and modified for use here by the statistical consultant. The values of T_a and T_b are found by inspecting the intercepts of the regression equations for each method.

$S_{a_1-a_2}$ = standard error of the differences

$$S^2_{a_1-a_2} = s_p^2\left({}^1/_{n_1} + {}^1/_{n_2} + \sum_{i=1}^{j} c_{1ii}\bar{x}_{1i} + 2\sum_{i=1}^{j}\sum_{m>i}^{j} c_{1im}\bar{x}_{1i}\bar{x}_{1m} + \sum_{i=1}^{k} c_{2ii}\bar{x}_{2i} + 2\sum_{i=1}^{k}\sum_{m>i}^{k} c_{2im}\bar{x}_{2i}\bar{x}_{2m}\right)$$

$$s_p^2 = \frac{SS_{1error} + SS_{2error}}{df_{1error} + df_{2error}}$$

Figure 2.2 Equations to Determine $S_{a_1-a_2}$

3. William G. Cochran, Planning and Analysis of Observational Studies, John Wiley & Sons, Inc., 1983.

It was here that the SAS GLM procedure was put to use again. In order to obtain the inverse matrix, C, it was necessary to first determine the matrix of the general model, $X'X$. The value of this matrix was determined by using the INVERSE option available with the GLM procedure. The last values that had to be ascertained were the means for the dependent variable and the independent variables in each equation. This was done by writing an auxiliary program using the programming language offered within SAS. With these values and the matrix, it was simply a matter of substituting them into the equations given in Figure 2.2 and looking up the results in the proper statistical table.

Chapter 3 - Results

The regression analysis on the entire data set produced two interesting observations. The first was that the variable associated with Method B did not enter the Stepwise Regression at any point during the execution of the procedure. The second observation made was that the variable tied to Method A was in the final regression equation produced by the Stepwise procedure.

Course	Variable	B Value
CMPSC305	A	13.9076
CMPSC405	A	8.7592
CMPSC460	A	9.9427

Table 3.1 b Values for Variable A

The effect of Method A was determined by inspecting the b values[1] for variable A given in Table 3.1. A positive value indicates a positive effect by that variable and a negative value vice-versa. Therefore, Method A had a positive effect on each of the three follow-up courses, as determined by this test.

A study of the values in Table 3.2 shows that for Method A, the intercept values are significant but that they are not significant for Method B. This deduction is best explained by the following passage from the SAS Users Guide;

> The value given by PR > |T| answers the question, "If the parameter is really equal to zero, what is the probability of getting a larger value of T?" Thus, a very small value for

1. The b values are the corresponding estimated regression coefficients for that variable in the regression equation.

this probability indicates that the parameter is not likely to equal zero, and therefore that the independent variable contributes significantly to the model."[2]

Course	Method	$T_{intercept}$	n	PR > |T|	SS_{error}	MS_{error}
CMPSC305	A	31.657	69	.0035	8172.74	125.73
CMPSC405	A	29.264	47	.0027	2741.71	62.31
CMPSC460	A	22.585	50	.0196	963.37	24.70
CMPSC305	B	13.317	63	.4962	8741.28	156.09
CMPSC403	B	2.495	39	.9142	2078.88	67.06
CMPSC460	B	-20.138	33	.0654	300.31	13.65

Table 3.2 Results from SAS Procedure GLM

Therefore, from these results, Method A again shows a positive effect, whereas, the results for Method B are not significantly different from zero.

Table 3.3 shows the results of computing the values for the equations given in Figure 2.2. Note that there are no values for CMPSC460. This is due to the negative value returned for the intercept on Method B. The values of importance in this table are those found in columns four and six.

course	s_p^2	$S_{a_1-a_2}$	t	df	t_{df}
CMPSC305	139.76	20.659	0.8877	121	1.980
CMPSC405	64.30	24.76	1.0812	75	1.992
CMPSC460	>>>>> values not computed <<<<<				

Table 3.3 Equation and t values

In order for the null hypothesis to be rejected, the value in column four has to be greater than the value in column six, but this is not so. The

2. SAS Users Guide, SAS Institute Inc., Cary, North Carolina, 1979, p 239.

observation to be made here is that the results of the Student t test showed no significance when the two methods were compared against each other.

Chapter 4 - Conclusions and Future Work

To reiterate, the purpose of this report has been to look at the effectiveness of two teaching methodologies that have been used in Algorithmic Processes at Kansas State University. To answer this, and other questions, this study was conducted on a group of undergraduate Computer Science majors who had completed Algorithmic Processes under one of these two methods. A list of the observations made, based on the statistical tests, are given below.

1) The regression analysis tests showed that Method A had a positive effect on each of the three follow-up courses.

2) The absence of Method B from the regression analysis is interpreted as meaning that this method had no effect on any of the follow-up courses.

3) The results of the GLM procedure indicate that the effect of Method A is significant while that of Method B is not.

4) The results of the direct comparison of the two methods were not significant, therefore no conclusion may be made.

Based upon these observations, it is apparent that Method A (Essentialism) does have an advantage over Method B (Information Processing) in it effectiveness of preparing students for study in each of the three follow-up courses.

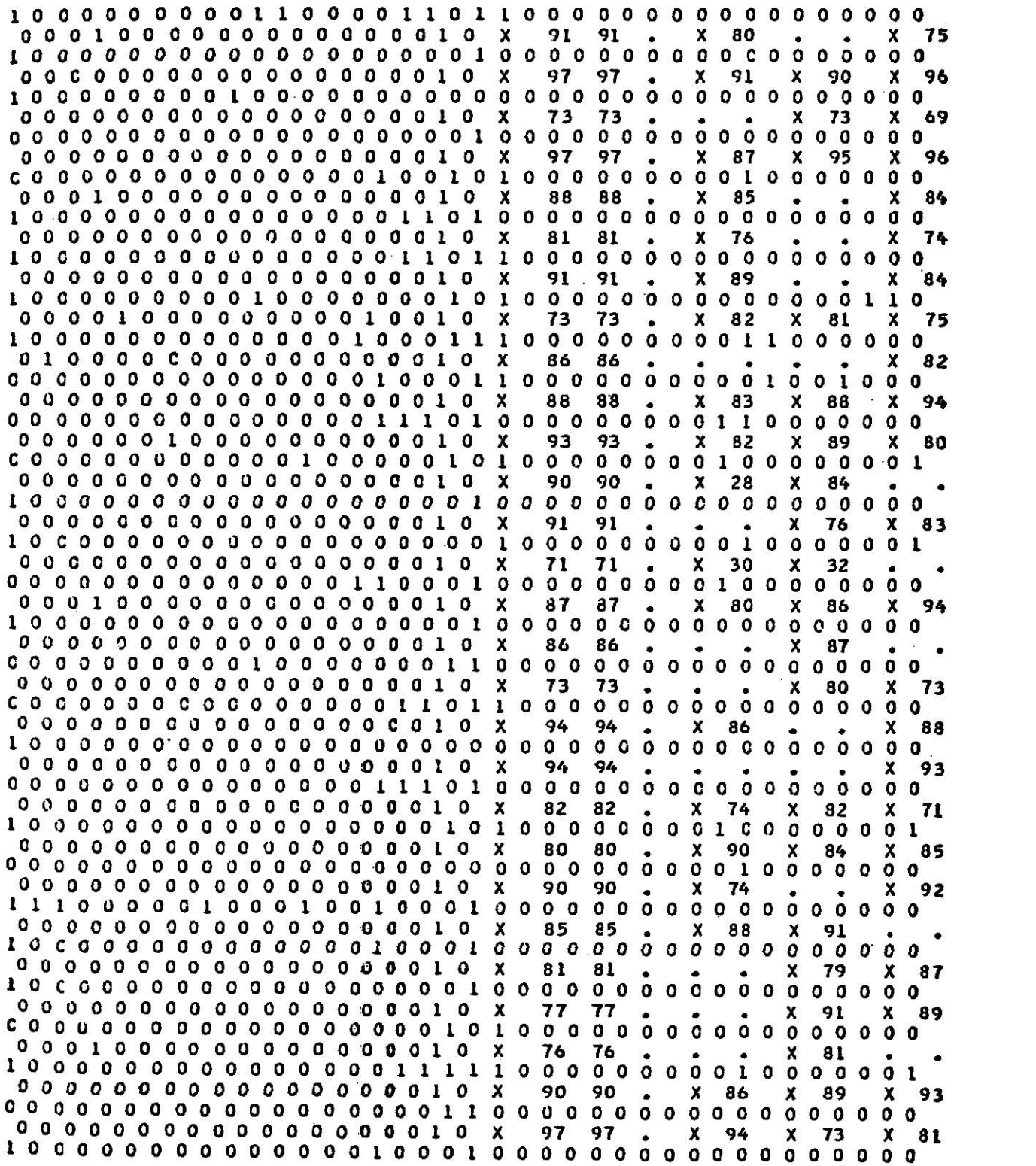
One possible explanation for the results of the comparative test is that the number of variables in the regression equation caused the error values for the sum of the squares to be too large. (see Table 3.2) As a

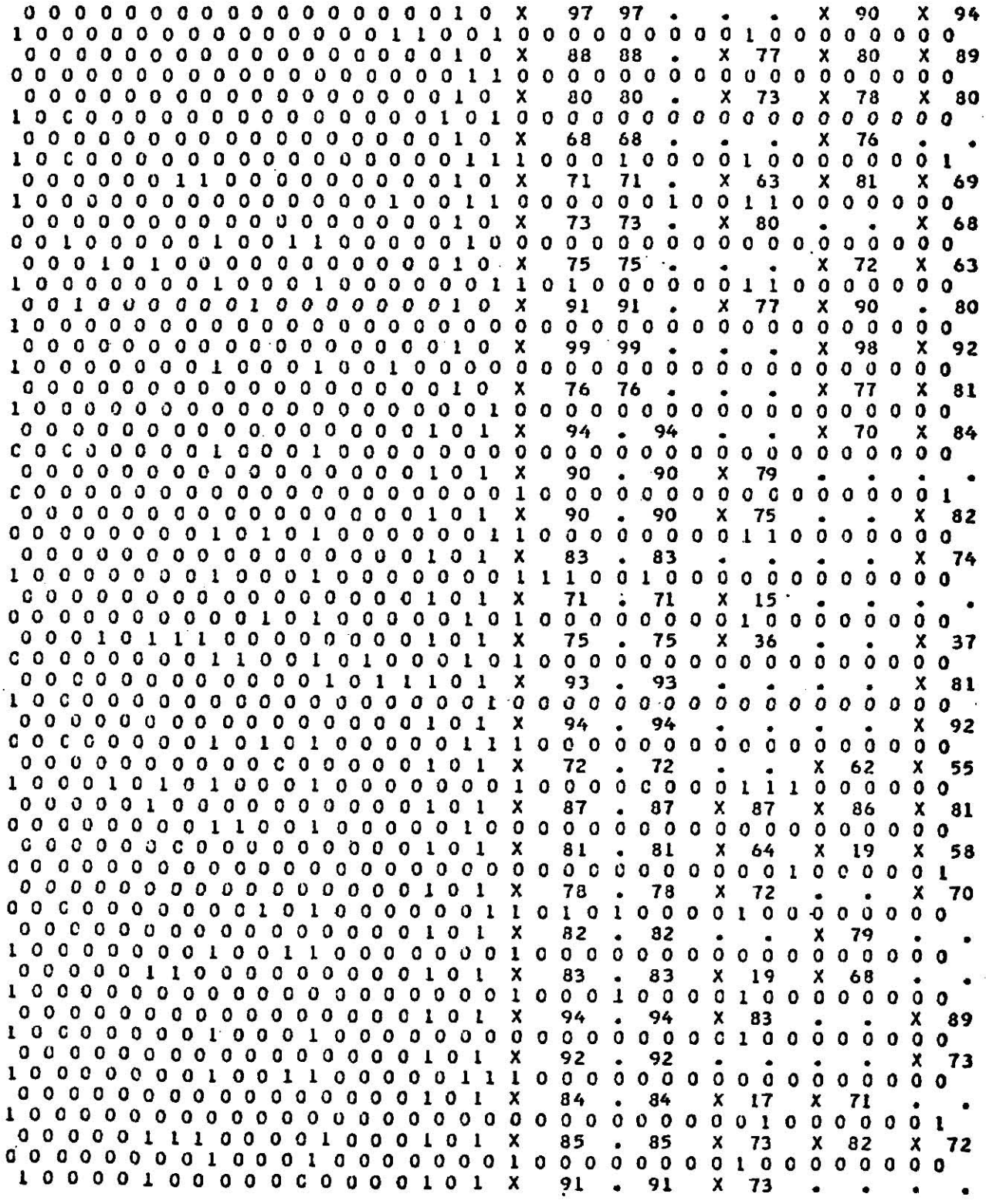
result, the test was not sensitive enough to the differences in the effects of the two methods.

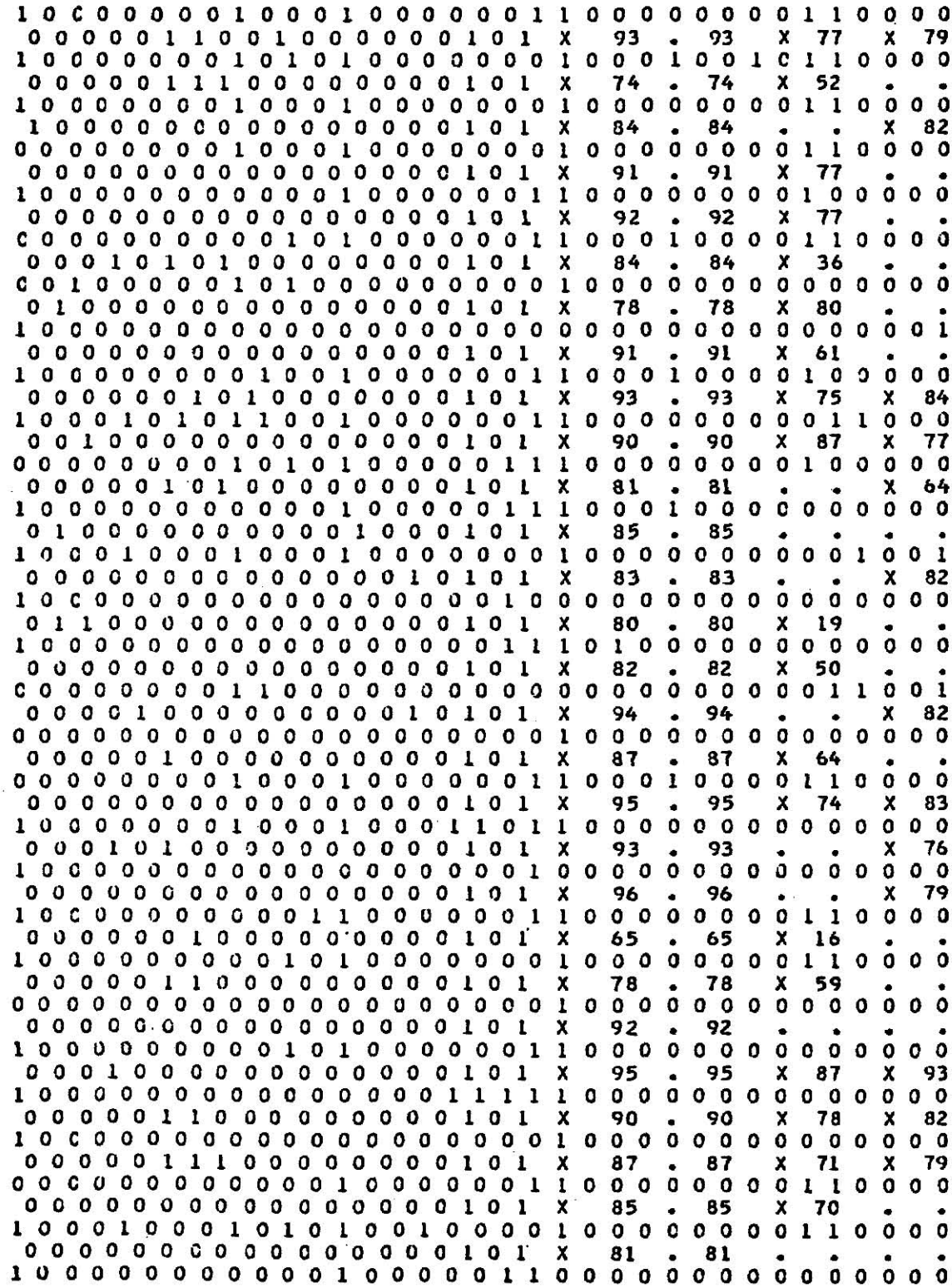
Future studies into this area would benefit from collecting more background information at the start of the study on each student. This would include the community size, high school size, high school GPA, high school academic background, SAT/ACT scores, and others deemed necessary. This could be accomplished by conducting a survey of each student upon enrollment in Algorithmic Processes and then tracking the progress of each student through all of the follow-up courses. A second method of conducting this study would be to have one instructor conduct two sections of Algorithmic Processes using one of the methods in each class. The enrollment in each section would have to be as random as possible. By doing this, the personality of the instructor would be removed from the results of the study.

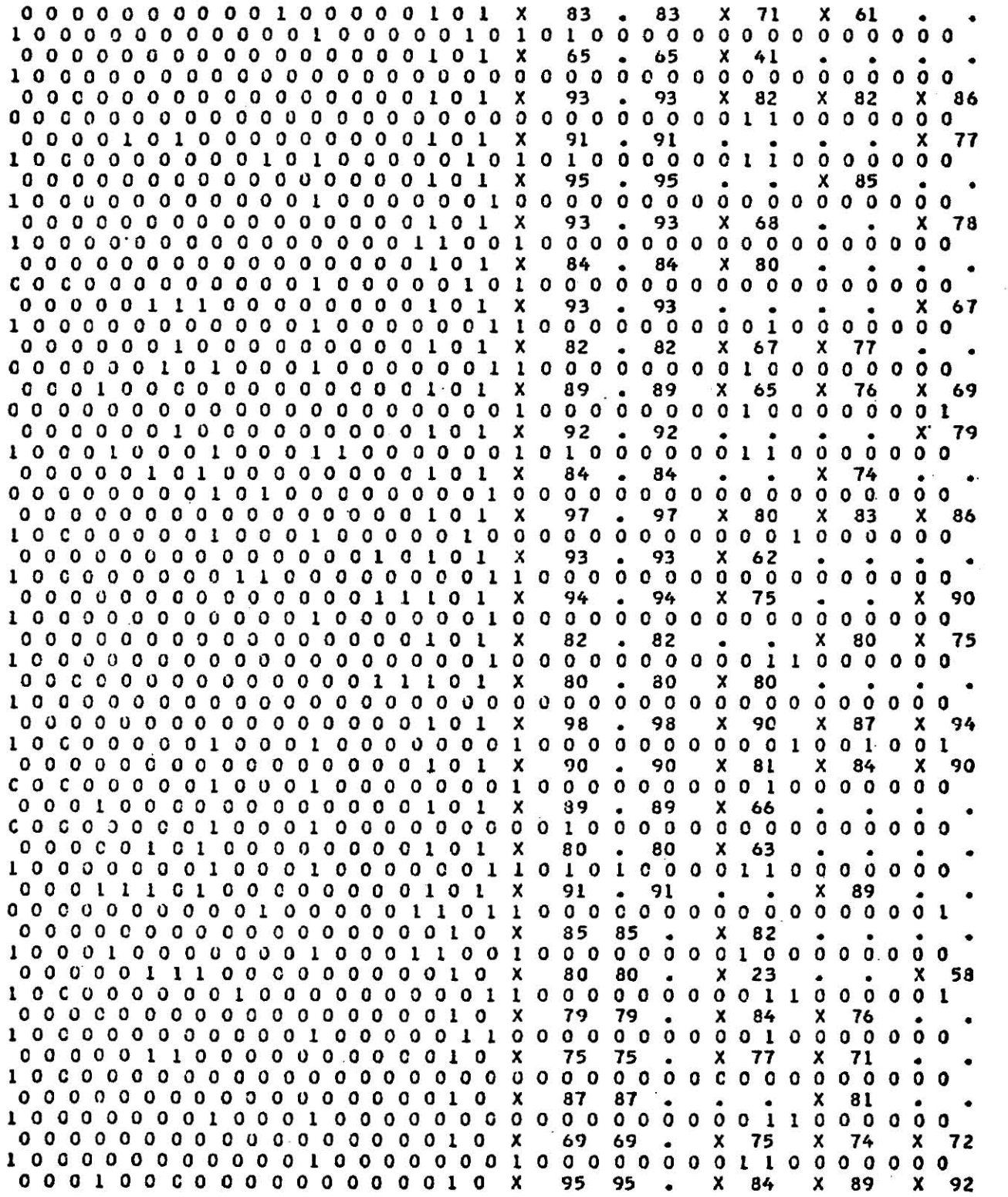
Appendix A

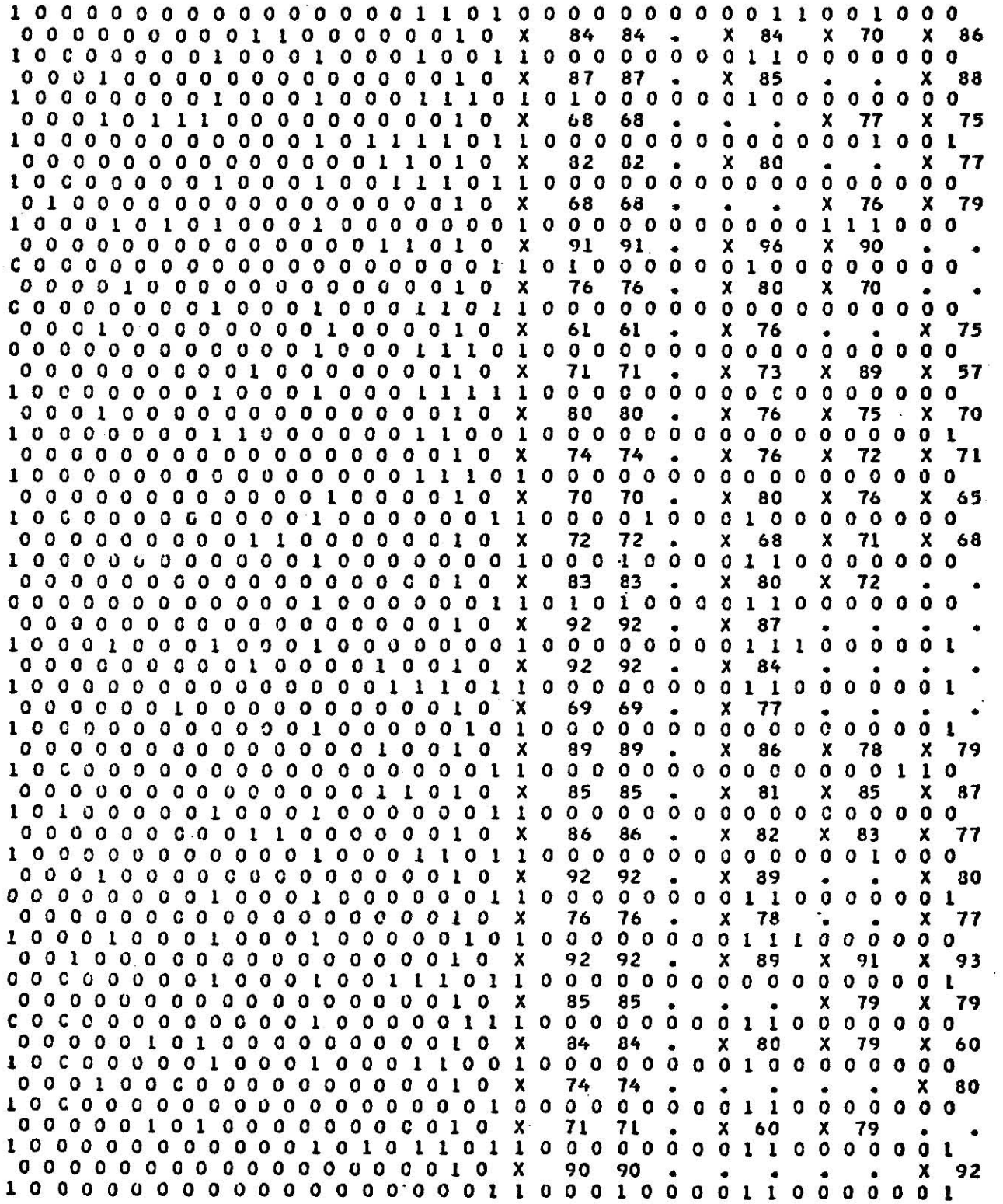
STUDENT OBSERVATIONS

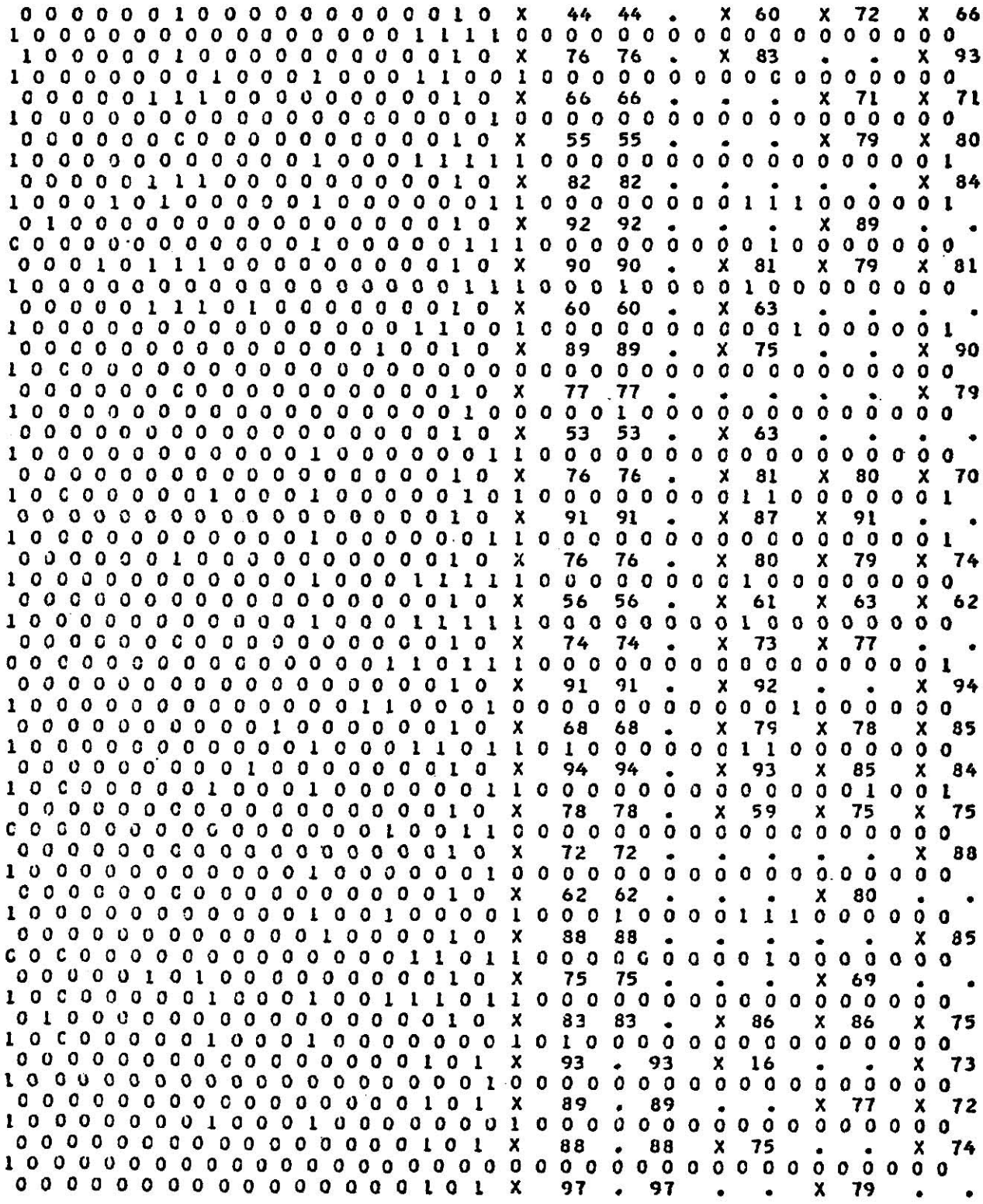


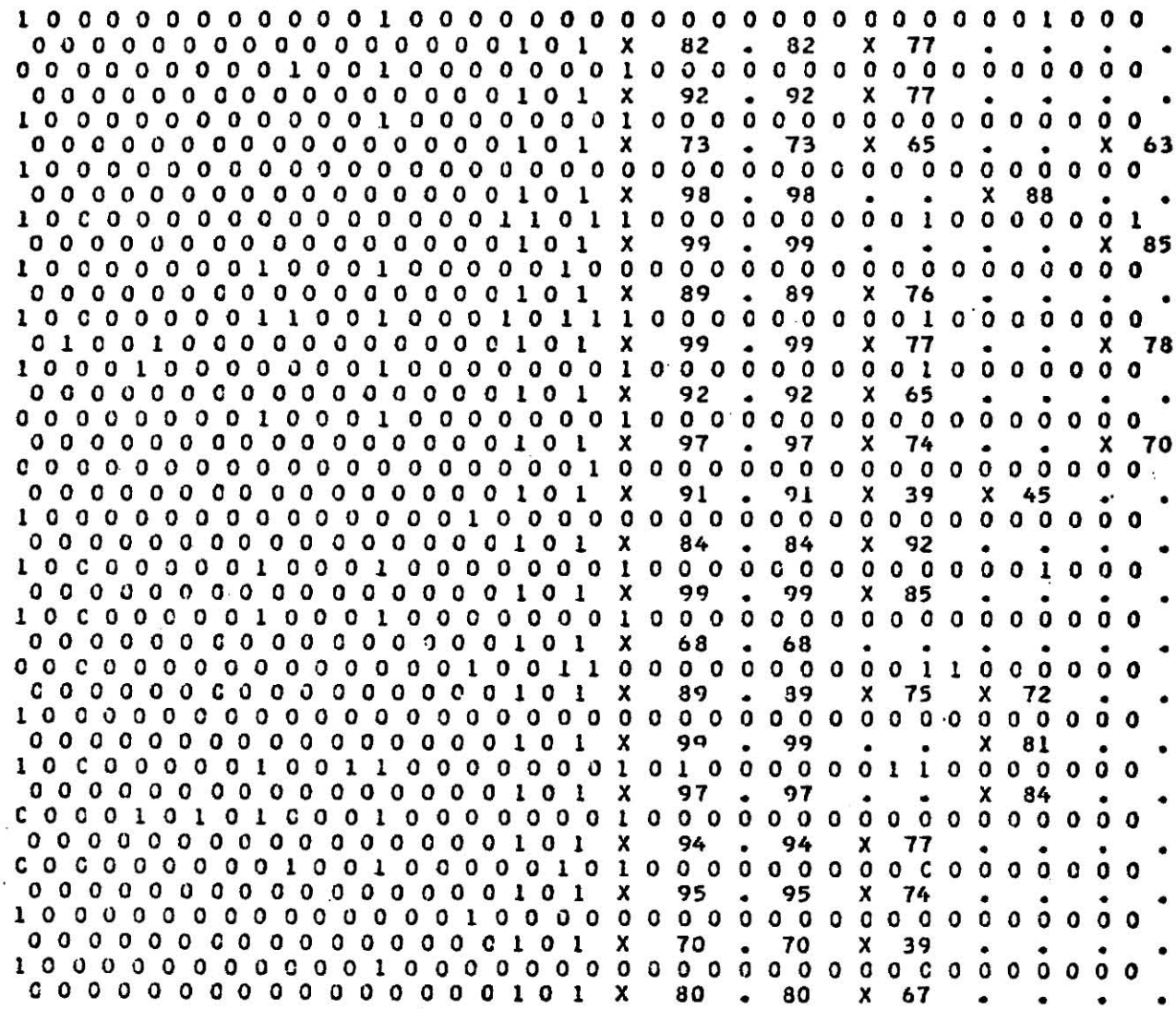












Appendix B

INDEPENDENT VARIABLES

Listed according to their entry numbers found in Table 2.3 in Chapter 2.

variable #	course
2	Biochemistry 110
3	Chemistry 110
4	Chemistry 190
5	Chemistry 210
6	Chemistry 220
7	Chemistry 230
8	Chemistry 271
9	Computer Science 200
10	Computer Science 201
11	Computer Science 202
12	Computer Science 206
13	Computer Science 207
14	Computer Science 211
15	Computer Science 306
16	Computer Science 307
17	Computer Science 340
18	Computer Science 341
19	Computer Science 362
20	Computer Science 420
21	Electrical Engineering 241
22	Mathematics 010
23	Mathematics 100
24	Mathematics 110
25	Mathematics 150
26	Mathematics 170
27	Mathematics 205
28	Mathematics 210
29	Mathematics 211
30	Mathematics 220
31	Mathematics 221
32	Mathematics 222
33	Mathematics 225
34	Mathematics 226
35	Mathematics 240
36	Mathematics 250
37	Mathematics 251

38	Philosophy 100
39	Philosophy 105
40	Philosophy 110
41	Philosophy 125
42	Philosophy 130
43	Philosophy 220
44	Physics 101
45	Physics 102
46	Physics 103
47	Physics 104
48	Physics 113
49	Physics 114
50	Physics 115
51	Physics 191
52	Physics 193
53	Physics 213
54	Physics 214

Appendix C

SELECTED REGRESSION EQUATIONS

Each set of equations are labeled according to where they were used.

Regression Analysis

Course	Model equation
CMPSC305	S305 = SEX X1-X53 A B S300
CMPSC405	S405 = SEX X1-X53 A B S300
CMPSC460	S460 = SEX X1-X53 A B S300

Analysis using the SAS GLM (General Linear Model) Procedure

Course	Method	Model equation
CMPSC305	A	S305 = X19 X45 S300
CMPSC405	A	S405 = S300 S305
CMPSC460	A	S460 = SEX X9 X12 X18 X31 X41 X47 X50 S300 S405
CMPSC305	B	S305 = X6 X11 X22 X40 X41 S300
CMPSC405	B	S405 = SEX X10 X15 X18 X20 X42 S300
CMPSC460	B	S460 = X2 X8 X17 X18 X22 X30 X34 X44 S300 X305

Bibliography

Bailey, Gerald D., Unpublished article on evaluation of teaching methodologies, College of Education, Kansas State University, 1983.

Bailey, Gerald D., "An Evaluator's Guide to Diagnosing and Analyzing Teaching Styles," National Association of Secondary School Principals, vol. 68, no. 469, February 1984, pp. 19-25.

Brown, George, Lecturing and Explaining, London: Methuen and Co., 1978.

Cochran, William G., Planning and Analysis of Observational Studies, John Wiley & Sons, Inc., 1983.

Draper, Norman and Harry Smith, Applied Regression Analysis, John Wiley & Sons, Inc., 1966.

Duck, Lloyd, Teaching with Charisma, Boston: Allyn and Bacon, 1981.

Eggen, Paul D., Donald P. Kauchak and Robert Harder, Strategies for Teachers -- Information Processing Models in the Classroom, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1979.

"A Nation at Risk: The Imperative for Educational Reform," The National Commission on Excellence in Education, U.S. Department of Education, April 1983.

SAS Institute, Cary, North Carolina (release 82.3 at Kansas State University).

A STUDY OF PEDAGOGICAL APPROACHES TO TEACHING PROBLEM SOLVING

by

BRIAN LYN SNYDER

B. A., Bethany College, 1977

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1985

ABSTRACT

This paper briefly outlines two methods of instruction which have been used in teaching Algorimthic Processes at Kansas State University. Through the use of selected statistical tests, an analysis of the effectiveness of each method was made. A description of the process of data collection is included along with a listing of the final data set. This set of observations was then tested using selected statistical tests to determine if either teaching method was more effective than the other. The results of these tests, possible improvements on the statistical tests, and areas for future study are also discussed.