DOCUMENT RESUME

ED 095 001 SE 017 724

AUTHOR	Aikenhead, Glen S.
TITLE	On Using Qualitative Data to Evaluate Two Chemistry Courses.
PUB DATE	Apr 74
NOTE	48p.; Paper presented at the annual meeting of the National Association for Research in Science Teaching (47th, Chicago, Illinois, April 1974)
EDRS PRICE	MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS	*Chemistry; Concept Formation; *Educational Research; *Evaluation; Grade 11; Grade 12; Science Education; Scientific Enterprise; Scientists; *Secondary School Science
IDENTIFIERS	Saskatchewan; *Science Processes

ABSTRACT

The purposes of this study were to extend the investigation concerning the limitations of quantitative data, explore the need for and use of qualitative data, and to evaluate two chemistry courses (CHEM Study and Modern Chemistry, 2nd edition). In his research, the investigator conducted a partial evaluation of these two courses and assessed the knowledge about science and scientists possessed by grade eleven and grade twelve students in Saskatchewan. Thirty-eight teachers were randomly selected to participate in the study; 28 agreed to do so. Student data were collected via the Nature of Science Scale--form SASK (NOSS), an instrument originally designed to assess the effect of university science programs on university students' understanding of the nature of science; the Science Process Inventory form D-SASK (SPI); and the Test on the Social Aspects of Science (TSAS) which deals with the interaction among science, technology and society. The use of t tests for matched pairs indicated that both chemistry courses had a negligible effect on students' knowledge about the scientific enterprise. Qualitative data revealed a balance between a gain in understanding and an acquisition of misconceptions. Also identified were specific ideas which indicated strengths and weaknesses in students' understanding about science and scientists. (PEB)

ON USING QUALITATIVE DATA TO EVALUATE TWO CHEMISTRY COURSES

Glen S. Aikenhead
Associate Professor
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

Paper presented at the 47th annual meeting of
the National Association for Research in Science Teaching.

CHICAGO

April 16-18, 1974.

TABLE OF CONTENTS

ON USING QUALITATIVE DATA TO EVALUATE TWO CHEMISTRY COURSES

Glen S. Aikenhead
Associate Professor
University of Saskatchewan
Saskatoon, Sask., Canada

INTRODUCTION

Quantitative data have specific limitations as far as giving feedback to teachers and curriculum developers is concerned. A complementary or alternative form of evaluation information could enhance the quality of the evaluation process. This alternative form of information emerges from asking such qualitative questions as: What ideas have students learned? What misconceptions did they acquire? What misunderstandings have they still retained?

At last year's NARST annual conference, a paper was presented (Aikenhead, 1973a) concerning: (1) the limitations of quantitative data, and (2) an exploration of the need for, and the use of, qualitative data. Project Physics was partially evaluated in order to illustrate the use and limitations of qualitative data.

The purposes of this study are: (1) to extend the investigation reported last year to NARST, and (2) to evaluate two chemistry courses (CHEM Study and Modern Chemistry -- 2nd edition). This evaluation information was used by the Saskatchewan Chemistry Curriculum Committee in its formative evaluation of its chemistry program.

THE PROBLEM

The general state of understanding science and the scientist would seem to be unchanged since the 1950's (Aikenhead, 1972, pp. 1-2, 23-26). However, it is not enough to bemoan the fact that the national science curricula of the 1960's have had inconsequential effects on an adolescent's realistic understanding of the scientific enterprise. Science educators need:

(1) a list of specific misconceptions that are generally held by a representative sample of students; and

(2) some ideas as to the impact science classes are now having with respect to this type of knowledge.

Only then can curricula be developed or modified to overcome present deficits in students' understanding.

In Saskatchewan, chemistry students study either CHEM Study (Freeman, Prentice-Hall, or Heath editions) or Modern Chemistry (Metcalfe, Williams & Castka; 2nd edition) in grades XI and XII. Approximately 65% and 51% of the grades XI and XII students enrol in chemistry, respectively.

This investigation:

(1) is a partial evaluation[1] of CHEM Study and Modern Chemistry taught in Saskatchewan in the 1972-73 school year; and

(2) assesses the knowledge about science and scientists possessed by grade XI and XII Saskatchewan chemistry students.

---

[1] Because student learning is defined by the three evaluative instruments used to collect the data, the evaluation is limited to students' knowledge about science and scientists (the content of the three instruments). In this sense, the study is a partial evaluation.

PROJECT DESIGN

Sample

From the Saskatchewan Department of Education's 1971-72 list of chemistry teachers (353 total), 38 teachers were randomly selected to participate in the study. Each teacher was contacted by letter and telephone in September 1972 and 37 agreed to have their chemistry students respond to some evaluation instruments. It is assumed that the students of these randomly selected teachers represent the population of Saskatchewan chemistry students, even though the students themselves were not randomly selected. The sample size was reduced to 28 when five teachers did not return the pretest responses, three were unable to have their students respond to the posttests, and one teacher did not teach chemistry as anticipated. (This reduced the number in the urban grade XI CHEM Study subsample more than any other group.) The resultant subgroup of 28 teachers randomly selected from a population of 353 teachers allows one to be fairly confident in generalizing one's results to the total population. (The similarity between the student sample in this study and the population from which the sample was picked is shown in Appendix A.)

Table 1 shows the distribution of teachers and students among two courses, two grades, and two administrative systems (studying the course during one semester or over a full year period). Because there were seventeen teachers who taught both grades XI and XII, one may think in terms of there being a total of 45 teachers, for the purposes of analyzing student responses.

Table 1 fits about here

TABLE 1

SAMPLE POPULATIONS (TEACHERS & STUDENTS)

		Grade XI				Grade XII			
		CHEM Study		Modern Chemistry		CHEM Study		Modern Chemistry	
	total	semester	full year	semester	full year	semester	full year	semester	full year
# of teachers	45	5	6	1	9	10	5	1	8
		total = 11		total = 10		total = 15		total = 9	
# of students	734	116	122	21	101	213	88	18	55
		total = 238		total = 122		total = 301		total = 73	

The number of students was originally 790, but this number was reduced by 7% because of absenteeism for the posttest. Only the students who wrote both the pretest and posttest were analyzed. A further 0.8% of the students were deleted due to their answer sheets being unusable.

Instruments

Three instruments were used to assess the students' knowledge about the scientific enterprise. The tests were evenly distributed among the students of each teacher. By this method, a maximum of information was obtained in a minimum of testing time (Walberg & Welch, 1967), however, the number of students in each subgroup writing one particular test was diminished by one-third.

Students were asked to respond to every test item on a five point scale: strongly agree, agree, I do not understand, disagree, strongly disagree. The inclusion of the response, "I do not understand," allowed a student to indicate that a choice was impossible due to vocabulary or syntax problems. This five-choice response format deviates from the suggested response format of the original instruments. The reasons for making such a change are discussed below.

A detailed description of the original instruments, including their validation and reliability data, may be found elsewhere (Aikenhead, 1973b). Only a cursory description is presented here.

The Nature of Science Scale -- form SASK (NOSS) was developed at Stanford to assess the effect of the university's science programs on a student's understanding of the nature of science (Kimball, 1965). The instrument has never been validated for a high school population. Therefore on the basis of a preliminary study (with second year nonscience university students), the difficult vocabulary was simplified and one item was dropped altogether. The Saskatchewan chemistry students' responses to the NOSS form SASK should give valuable data concerning the instrument's suitability for high school students. The NOSS has 28 items.

The Science Process Inventory form D-SASK (SPI) deals with a student's awareness of the activities, assumptions, products and ethics of science (Welch, 1966). The form of the SPI usually used by researchers, form D, has a forced-choice response format--students are forced to agree or disagree with each item. In a former investigation (evaluating Project Physics), it was found that the forced-choice response format led to ambiguities for the researcher. One did not know if a student was disagreeing with a statement because he really did disagree or whether he disagreed because he did not understand what the statement meant (Aikenhead, 1972, p. 221). The same ambiguity exists for an agree response. Therefore, a Likert type scale (described above) replaced the agree/disagree format of form D. Because students are not forced to guess with the Likert type scale, the average scores should drop from form D to form D-SASK. (Comparing Saskatchewan students with United States norms was not an objective of this study.) The SPI has 135 items.

The Test on the Social Aspects of Science (TSAS) deals with the interaction among science, technology and society; the social nature of the scientific enterprise; and the social and political responsibilities of scientists (Korth, 1968). Except for one slight alteration, the original response format is the same as the one used in this study. "I do not understand" replaced "uncertain or undecided." The change was made in order to force a student to take a stand if he thought he understood the meaning of the item. The TSAS has 52 items.

All three tests were marked in the same manner. Each item has a "correct" response based on the content validity of the instrument. The response "agree" or "strongly agree" or the response "disagree" or "strongly disagree" was considered correct if it concurred with the test's key. The "I do not understand" response was always marked incorrect, however it was not ignored. Each item was later analyzed with respect to the percentage of students who agreed with, did not understand, and disagreed with, the item.

Many of the NOSS's items appear (verbatum or in modified form) in the SPI or the TSAS. Therefore, one could say that a high degree of content overlap exists between the SPI and the NOSS and between the TSAS and the NOSS.

Table 2 shows the tests' quantitative attributes based on the performance of the students in this study. These data are not intended for a comparison of Saskatchewan students with American norms. As mentioned above, the change in format for the SPI pre-empts such a comparison, and data for the NOSS have not been reported for high school students. It is of passing interest, however, that the Saskatchewan means and standard deviations for the TSAS are not unlike those of Korth's survey group (28.30 points and 4.98; Gallagher, 1969).

Table 2 fits about here

The reliability of the three tests is not exactly comparable because each test has a different number of items. Ferguson (1966, p. 381) suggests a way to transform reliability coefficients (r_{xx}) into a theoretical value (r_{kk}) based on an 100 item test:

$$r_{kk} = \frac{k \cdot r_{xx}}{1 + (k - 1)r_{xx}}$$

where: k = 100/(# of items in the test)

TABLE 2

QUANTITATIVE DATA FOR THE NOSS (form SASK), SPI (form D-SASK), AND TSAS BASED ON SASKATCHEWAN CHEMISTRY STUDENTS IN GRADES XI AND XII

			Mean Score	SD[a]	Range	Pre - Post Correlation	Reliability[b] estimate	N
NOSS:	Grade XI	pretest	9.00	2.436	3-15		.15	
						.33		122
		postest	3.87	2.532	3-18		.21	
	Grade XII	pretest	9.68	2.854	4-16		.38	
						.47		133
		postest	7.35	2.829	4-17		.38	
SPI:	Grade XI	pretest	92.09	12.500	63-120		.85	
						.63		107
		postest	94.50	12.780	65-117		.86	
	Grade XII	pretest	98.29	10.606	61-122		.81	
						.71		112
		postest	99.47	11.301	57-118		.84	
TSAS:	Grade XI	pretest	28.41	5.163	15-41		.63	
						.62		131
		postest	28.63	5.532	17-42		.67	
	Grade XII	pretest	30.64	4.878	18-43		.50	
						.62		129
		postest	30.84	5.376	15-42		.67	

[a] SD means standard deviation

[b] Kuder Richardson formula -20 was used.

For example, the theoretical reliability coefficients (based on an extrapolated 100 item test) for the SPI, NOSS, and TSAS are: .82, .69, and .80 respectively.

The reliability of the NOSS is much lower than most acceptable standards. Together with the fact that there is a high degree of content overlap between the NOSS and the other two tests, this indicated that in the future studies one could readily ignore the NOSS and use just the SPI and TSAS.

QUANTITATIVE RESULTS

Student responses to the three instruments were machine scored and then analyzed[1] both quantitatively and qualitatively. The quantitative analysis is described first.

The students were grouped according to their year of study (grades XI and XII) and the course studied (CHEM Study and Modern Chemistry). A further dichotomy within each group (semester and full year administrative systems) was not feasible due to the limited number of students in the sample. However, this comparison was made by regrouping the data into the semester/full year dichotomy and by analyzing the responses a second time. The results of this analysis follow the present discussion. Tables 3-6 show the student achievement based on total test scores. Diagram 1 summarizes these four tables.

Tables 3-6 fit about here

[1] Appropriate computer programs were used for the data analysis at the University of Saskatchewan's Computer Centre.

TABLE 3

TESTS FOR SIGNIFICANCE BETWEEN PRETESTS AND POSTESTS FOR GRADE XI CHEM STUDY STUDENTS

		Mean	Gain	t test[a]	p<[b]	SD[c]
NOSS (N = 39)	pretest	9.10	-0.23	0.70	ns[d]	2.534
	postest	8.87				2.578
SPI (N = 70)	pretest	93.54	1.37	1.01	ns	11.998
	postest	94.91				12.605
TSAS (N = 89)	pretest	28.51	0.37	0.71	ns	5.416
	postest	28.88				5.641

[a]test for significance of the difference between two means for correlated samples
[b]probability of the t test value occurring by chance alone (cut off value of p is .05)
[c]standard deviation
[d]not significant at the .05 level of probability

TABLE 4

TESTS FOR SIGNIFICANCE BETWEEN PRETESTS AND POSTESTS FOR GRADE XI MODERN CHEMISTRY STUDENTS

		Mean	Gain	t test	p<	SD
NOSS (N = 43)	pretest	8.81	0.05	0.11	ns	2.234
	postest	8.86				2.237
SPI (N = 37)	pretest	89.35	4.35	2.69	.02	12.962
	postest	93.70				13.069
TSAS (N = 42)	pretest	28.21	-0.10	-0.15	ns	4.575
	postest	28.12				5.256

TABLE 5

TEST FOR SIGNIFICANCE BETWEEN PRETESTS AND POSTESTS FOR GRADE XII CHEM STUDY STUDENTS

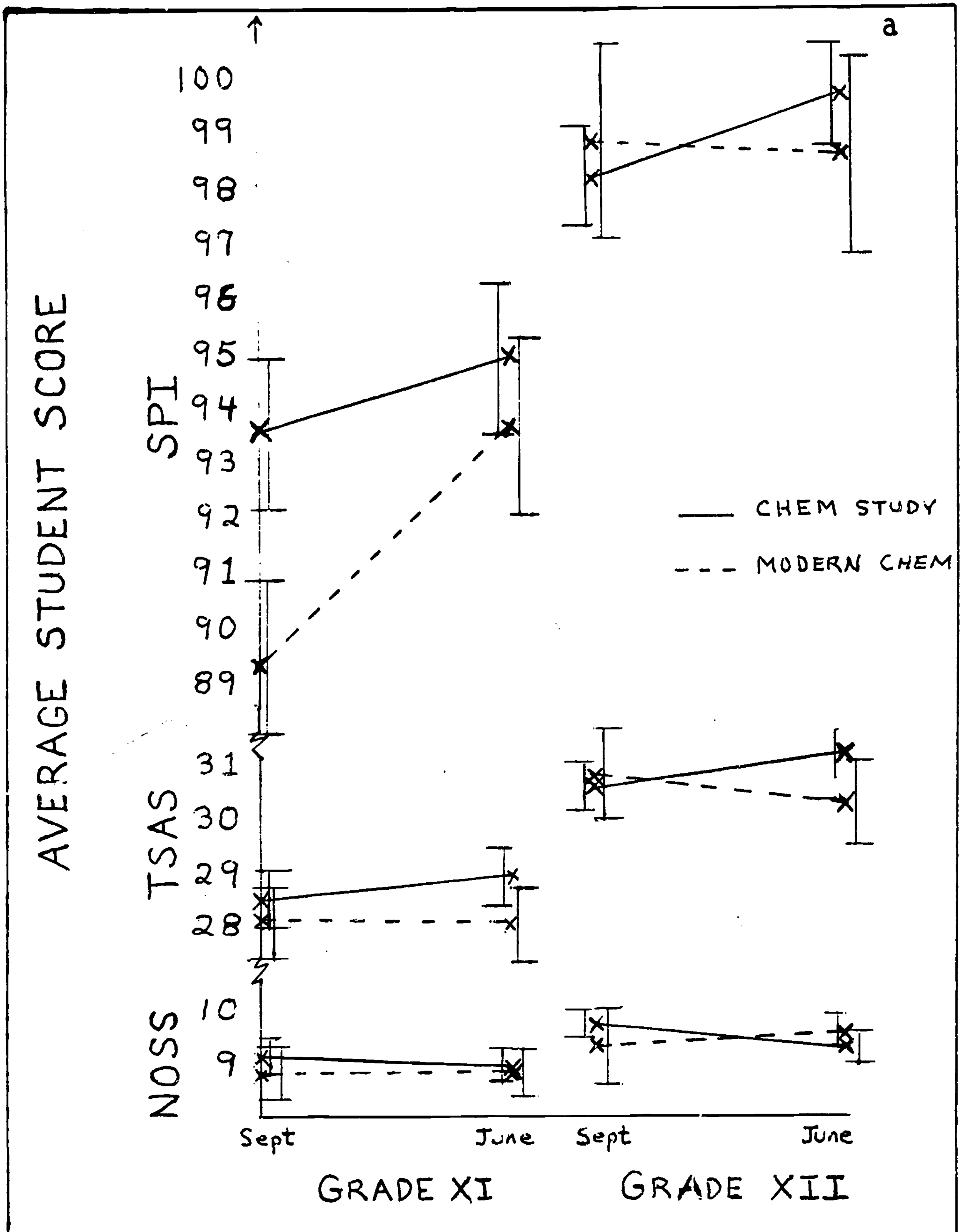
		Mean	Gain	t test	p <	SD
NOSS (N = 108)	pretest	9.76	-0.44	-1.65	ns	2.815
	postest	9.31				2.840
SPI (N = 91)	pretest	98.16	1.51	1.70	ns	11.039
	postest	99.67				11.388
TSAS (N = 102)	pretest	30.61	0.40	0.88	ns	4.806
	postest	31.01				5.386

TABLE 6

TEST FOR SIGNIFICANCE BETWEEN PRETESTS AND POSTESTS FOR GRADE XII MODERN CHEMISTRY STUDENTS

		Mean	Gain	t test	p<	SD
NOSS (N = 25)	pretest	9.32	0.20	0.28	ns	2.990
	postest	9.52				2.773
SPI (N = 21)	pretest	98.81	-0.19	-0.11	ns	8.461
	postest	98.62				10.874
TSAS (N = 27)	pretest	30.74	-0.52	-0.66	ns	5.139
	postest	30.22				5.294

FIGURE 1



[a] A vertical bar defines a critical region surrounding an average score. If one critical region overlaps with another critical region, then there is no significant difference between the two scores (at the .05 probability level).

On the basis of total test scores, only the grade XI Modern Chemistry group showed any significant gains, 4.35 points on the SPI. The same group made no significant gains on the other two tests. Although 4.35 points may represent a statistically significant improvement (approximately a one-third standard deviation), one could still be disenchanted with its being a fairly small gain for a test which has 135 items.

The grades XI and XII CHEM Study group and the grade XII Modern Chemistry group made no significant gain in total test scores.

Figure 1 fits about here

When all the data were dichotomized between the semester and full year administrative systems (pooling grade XI and XII along with CHEM Study and Modern Chemistry students) significant changes in total test scores resulted for the full year group on the SPI only (pretest 92.29, posttest 95.75, gain 3.46, for 101 students).

QUALITATIVE RESULTS

Although the data analyzed above yield quantified statements, these results are ambiguous. For example, what does a 4.35 point gain on the SPI really mean? This ambiguity leads one to ask qualitative questions such as: What ideas have students learned? What misconceptions did they pick up? What misunderstandings have they still retained? These questions are answered by two kinds of item analyses.

First of all, each item of each test is examined to see if students' correct responses significantly gained or dropped over the year. This analysis reflects the impact of each chemistry course on student learning. Every item is analyzed with respect to the changes in student responses between the pretest and posttest. There are only two changes possible: (a) from an incorrect response to correct response, and (b) from a correct response to an incorrect response. The probability is .50 that students who change their response move to a correct answer. McNemar (1969) has derived a chi square test specifically for this situation:

$$\chi^2 = \frac{(A - D)^2}{(A + D)}$$

with 1 degree of freedom[1], where A and D are cell frequencies in the following contingency table:

Item X

		Posttest 1	Posttest 0
Pretest	0	A	B
	1	C	D

1 signifies a correct response

0 signifies an incorrect response

The chi square analysis determines which items experience a statistically significant change in student response between the pretest and posttest.

Secondly, the percentage of student agreement, not understanding, and disagreement is also analyzed for each item. The results of this analysis are discussed below in the sections: "An Analysis of Student Correct Responses," and "An Analysis of Student 'I Do Not Understand' Responses."

---

[1] When $A+D < 20$, a Yate's correction for small frequencies must be introduced. The equation becomes: $\chi^2 = (|A-D| - 1)^2 / (A+D)$.

The Impact of the Chemistry Courses

Tables 7-14 list the items which experienced significant gains or losses. Except for the grade XI Modern Chemistry group, students generally gained as much misinformation as they did an understanding of the scientific enterprise. (This result was also reflected in the lack of significant gains or losses in total test scores, Tables 3-6.)

As shown in Tables 7 and 8, the grade XI CHEM Study students tended to learn something about: (1) scientific theories, (2) the interaction of science and society, and (3) the interdependence of science and technology. They also learned a little concerning limitations to science. Meanwhile some misconceptions within the CHEM Study group worsened. Students in both grades XI and XII tended to learn that there is no room for imagination or intuition when doing science, and that there are natural phenomena which should not be investigated by scientists. The 11th graders became more confused about: (1) the aims of science and the aims of technology, (2) the nature of quantitative data, and (3) the use of the simplicity value in scientific thinking.

Tables 7 & 8 fit about here

Oddly enough the grade XII CHEM Study students increased their understanding of the aims of science and technology (recovering from the "mislearning" of grade XI). As illustrated in Tables 9 and 10, these students became more aware of the interaction of science with society beyond the technological interface. They also tended to learn something about hypotheses and deduction. However, the 12th graders picked up some

TABLE 7

GRADE XI CHEM STUDY SIGNIFICANT GAINS [a]

	item	correct response	% pre	% post
NOSS 13	A scientist can only do research well when he works within his very narrow specialty. Otherwise, a team of researchers from various special fields must work together.	disagree	19	33
SPI 8	The essential test of a scientific theory is its use in predicting future events.	agree	40	61
SPI 102	Some presently accepted theories were opposed by other scientists when first proposed.	agree	80	91
TSAS 2	Statements are not accepted as scientific knowledge unless they are absolutely true.	disagree	23	33
TSAS 8	It is extremely difficult to anticipate how new scientific knowledge may affect society.	agree	70	80
TSAS 23	The economic prosperity of most nations today depends on their ability to discover and use scientific knowledge.	agree	58	75
TSAS 26	In modern industrial societies science and technology have little to do with each other.	disagree	72	84
TSAS 52	Many of today's social and political problems are related to science and technology. Since scientists are experts in this field we should accept their judgment in such matters.	disagree	51	61

[a]Significant gain means that there is less than .05 probability that the gain occurred by chance alone.

TABLE 8

GRADE XI CHEM STUDY SIGNIFICANT LOSSES[a]

	item	correct response	% pre	% post
NOSS 1	The most important scientific ideas have been the result of a systematic process of logical thought	disagree	20	9
NOSS 14	If a scientist investigates the possibilities of creating life in the laboratory, his research is an invasion into areas where science doesn't belong.	disagree	63	48
SPI 9	If two different hypotheses fit the observed facts, the simpler is accepted.	agree	35	21
SPI 67	Scientists assume that all effects in nature have causes.	agree	96	81
SPI 91	A measurement expressed as 12 inches is a statement of approximation.	agree	73	57
SPI 98	A measurement of length expressed as 15 inches is a statement of exact truth.	disagree	79	65
TSAS 22	The principal aim of science is to provide the people of the world with the means for living better lives.	disagree	24	14
TSAS 28	The steam engine was one of the most important developments in the history of science.	disagree	31	20

[a]Significant loss means that there is less than .05 probability that the decrease occurred by chance alone.

erroneous ideas concerning a scientist. In addition to the distorted knowledge described in the paragraph above, students tended to increase their misunderstanding regarding: (1) the role of opinion and fact in making scientific conclusions, (2) what a scientific theory is, (3) the assumption of nature's susceptibility to human understanding, (4) the role of open communication, and (5) the practicality value applied to scientific knowledge.

Tables 9 and 10 fit about here

It was the grade XI Modern Chemistry group which made the significant gain in their total test scores. Table 11 shows that these students became less naive about: (1) "The Scientific Method," (2) models duplicating reality, and (3) the finality and inalterable truth of scientific knowledge. Besides learning more about the relationship between evidence and theory, the 11th graders came to see science as less complicated to understand. They also realized that scientists themselves do not assume that some problems are to complex to be explained. On the other hand, some misconceptions were learned. The students tended to learn that (1) scientists do not do library research (only high school students?), (2) serendipity plays no important role in science, and (3) theories may not be modified in light of new evidence (a paradox to their achievement described above).

Tables 11 & 12 fit about here

TABLE 10

GRADE XII CHEM STUDY SIGNIFICANT LOSSES

	item	correct response	% pre	% post
NOSS 14	If a scientist investigates the possibilities of creating life in the laboratory, his research is an invasion into areas where science doesn't belong.	disagree	69	55
SPI 41	Science must start with facts and end with facts no matter what theoretical structures it builds in between.	agree	50	37
SPI 54	"We are going to have 36 snowfalls this winter" is an example of a scientific theory.	disagree	74	63
SPI 60	Scientific conclusions should be based on facts, not opinion.	agree	95	88
SPI 65	Investigation of the possibilities of creating life in the laboratory is an invasion of science into areas where it doesn't belong.	disagree	62	51
SPI 74	Scientists assume that the human mind is capable of understanding the events and materials of nature.	agree	79	65
TSAS 15	A scientist is expected to share the knowledge he discovers with other scientists rather than use it exclusively for his own profit.	agree	89	81
TSAS 32	A great research scientist is usually little concerned with the practical applications of his work.	agree	21	13

TABLE 9

GRADE XII CHEM STUDY SIGNIFICANT GAINS

	item	correct response	% pre	% post
SPI 58	A hypothesis is a simple guess or "hunch" that tries to explain several observations.	agree	52	69
SPI 63	Scientists view events today as clues to events in the past.	agree	78	91
SPI 89	Deduction is the process of predicting particular occurences from the general case.	agree	52	67
SPI 105	A value of a hypothesis is its suggestion of new experiments.	agree	78	88
SPI 119	A scientist believes that an experiment performed today will produce the same results as the same experiment performed last week.	agree	46	57
TSAS 19	Although advances in science and technology may improve living conditions they offer little help in solving today's social problems.	disagree	29	44
TSAS 22	The principal aim of science is to provide the people of the world with the means for living better lives.	disagree	12	21
TSAS 26	In modern industrial societies science and technology have little to do with each other.	disagree	81	94
TSAS 47	New scientific knowledge affects society only through the practical use made of it.	disagree	31	42

TABLE 11

GRADE XI MODERN CHEMISTRY SIGNIFICANT GAINS

	item	correct response	% pre	% post
NOSS 3	Thanks to the discovery of the scientific method, new discoveries in science have begun to come faster.	disagree	2	12
SPI 28	The knowledge of science is final.	disagree	65	89
SPI 59	Scientific models are exact duplications of reality.	disagree	68	87
SPI 63	Scientists view events today as clues to events in the past.	agree	68	89
SPI 77	Scientists assume that some problems are too complex ever to be explained.	disagree	35	65
SPI 94	Prediction is an important goal of scientific investigation.	agree	63	78
SPI 110	A theory with ten supporting and two denying experiments is more likely to be accepted than a theory with four supporting and no denying experiments.	disagree	41	65
SPI 111	The law of conservation of energy is an example of an unchanging truth.	disagree	19	41
TSAS 16	Modern science is too complicated for the average citizen to understand and appreciate it.	disagree	27	43

TABLE 12

GRADE XI MODERN CHEMISTRY SIGNIFICANT LOSSES

	item	correct response	% pre	post
SPI 1	Surprising or unexpected observations have played an important role in the advance of science.	agree	100	89
SPI 25	A theory in science may be modified in light of new evidence.	agree	89	79
SPI 122	When confronted with a new problem, a scientist searches the literature to see what similar work has been done.	agree	95	78
TSAS 49	It would be a good idea to slow science until society has had a chance to adjust to the changes science has brought about.	disagree	69	31

TABLE 13

GRADE XII MODERN CHEMISTRY SIGNIFICANT GAINS

		item	correct response	% pre	% post
SPI	68	Scientists assume that if events A and B occur at the same time, then one must be the cause of the other.	disagree	48	76
SPI	114	When a scientist makes a prediction, he is assuming that nature is consistent.	agree	71	91

TABLE 14

GRADE XII MODERN CHEMISTRY SIGNIFICANT LOSSES

		item	correct response	% pre	% post
NOSS	9	While biologists use the deductive approach to a problem, physicists always work inductively.	disagree	36	8
SPI	62	Induction is the process of generalizing the characteristics of a class from observations of all of its members.	disagree	33	14

11E

The 12th grade Modern Chemistry students made the least significant gains or losses. As indicated by Tables 13 and 14, these students became more aware of the causality and consistency assumptions but tended to become misinformed concerning induction and deduction.

Tables 13 & 14 fit about here

In comparing the achievement of the four groups of students, one finds very little similarity in the items which underwent significant gains or losses. Thus, each course at each grade level would seem to represent different learning experiences.

It is also interesting to consider the actual number of items that showed a significant change in correct response between the pretest and post-test. The present investigation may be compared with a similar study which investigated Project Physics and other physics courses (Aikenhead, 1973a). Compared with American physics courses, the Saskatchewan chemistry courses led to very few significant gains and led to a large number of significant losses.

Any significant CHEM Study gains tended to relate to the social aspects of science, while any Modern Chemistry gains tended to relate to the processes of science.

An Analysis of Student Correct Responses

The Saskatchewan chemistry students did have a good understanding in certain areas of knowledge about the scientific enterprise. These areas may be defined by the content of the items for which more than 85% of the students responded correctly. (Of course, there may well be items so important that one would not be satisfied with anything less than 100% correct student response. This decision is left to the reader.) On the other end of the scale, there were definite, widespread misconceptions about science and scientists among the students. These may be revealed by the content of items for which less than 50% of the students responded correctly.

The "above 85%" and "below 50%" criteria are arbitrary. They led to the identification of 41 and 65 items, respectively. These two numbers are unbalanced due to the NOSS data. No items from the NOSS were above the 85% correct response mark, but 23 items were below the 50% correct response mark. The SPI and TSAS showed a balance in the number of items above the 85% and below the 50% correct response mark.

A complete list of items meeting the "above 85%" and the "below 50%" criteria would be unreasonably long to include in this report. However, Appendix B identifies these items and a descriptive summary is given below. This description of Saskatchewan chemistry students' strengths and weaknesses has been organized into several content areas: the aims of science, its epistemology, its tactics, its values, its interactions with society, and its human needs. This categorization scheme is purely a heuristic one (Aikenhead, 1972, Chapter 1). The purpose of presenting the data in such a way is to help Saskatchewan teachers: (1) to identify students' strengths and weaknesses, and (2) to consider possible changes which might be made to their chemistry classes.

For the purposes of this general description, all students are analyzed as one single group, Saskathcewan chemistry students, because there were only minor differences in the students' achievement: (1) between the grade XI and grade XII data (with a slight trend toward a higher achievement on the part of the 12th graders, and (2) between pretest and posttest data (except for those cases already discussed above in connection with Tables 7-14). In addition, the only items considered for discussion are those which meet the "above 85%" and "below 50%" criteria for three or four of the original four groups of students (CHEM Study grade XI and XII, and Modern Chemistry grades XI and XII).

Aims of Science

Most students recognized that scientists do formulate theories to explain observations. However, the aims of science were hopelessly confused with the aims of technology. The students did not understand: (1) that the principal aims of science are: (a) to explain natural phenomena in terms of principles and theories, and (b) to reduce observations to a collection of mathematical relationships; and (2) that curiosity and a desire for self-esteem primarily motivate scientists.

More often than not, students wrongly considered that the primary object of basic scientific research is: (1) to provide the people of the world with the means for leading better lives, and thus (2) to produce new and improved machinery and living conveniences. In summary, less than half of the students correctly agreed to the statement, "The main object of basic scientific reseach is the discovery of understanding rather than its practical applications."

Epistemology of Science

There are many facets to the epistemology of science including: definitions, assumptions, observations, hypotheses, theories, models, laws, and general aspects of scientific knowledge.

Definitions

Two-thirds of the students erroneously thought that science was simply "an organized body of knowledge." A majority of students were confused, or did not understand, the meaning of induction and deduction.

Assumptions

While most respondents believed scientists assume that "all effects in nature have causes," many incorrectly thought scientists assume that "some natural things will never be understood." Students did not know that scientists accept assumptions before "they are proven true," and a large number disagreed with the consistency assumption.

Observations

Students did very well in their understanding the nature and role of observations in science. Seven items belonged to the "above 85%" category while none were poorly done. Carefully recorded, accurate observations were seen as important features of modern science. Just because measurements involve numbers did not mean the measurements cannot be wrong. Students recognized degrees of preciseness in examples of measurements, and they realized that two people making the "same" observation may see different things.

Hypotheses and Theories

Most students understood that hypotheses "may be wrong," and that theories "may be modified in light of new evidence." But when given an incorrect example of a hypothesis, these students mistook it to be a good example. The majority of students also felt that hypotheses have more experimental support than theories, thereby revealing an inaccuracy in their definition of the terms.

Models

A major ambivalence emerged from students' ideas about scientific models. Regarding the Bohr model of the atom, over 85%of the students recognized it as a "man-made" convenient representation for the purpose of understanding the atom. On the other hand, these students carried with them the misconception that the Bohr model "pictures the atom as we actually know it to exist," because it is "a scaled-up picture of what scientists have seen in their microscopes." Another paradox was discovered. Over half of the students wrongly disagreed with the statement, "Many scientific models (like the model of the atom) are made up through man's imagination, they do not duplicate reality;" while over two-thirds correctly disagreed with the statement, "Scientific models are exact duplications of reality." Sifting through this data, one may conclude that students have superficial ideas concerning the nature of scientific models; models are indeed man-made but not by man's imagination.

Laws

Again it appears as if a superficial understanding exists about scientific laws. Most students correctly believed that when a statement becomes a law of science, it can still be changed. However, over half of the students affirmed, most incorrectly, that "a physical law is permanent" and that "the law of conservation of energy is an example of an unchanging truth." Perhaps the students believe that statements may change but laws do not -- a very peculiar state of mind. On the positive side, most students were not naive enough to believe that nature cannot disobey the laws of science.

General Aspects

On the one hand, many students asserted that scientific knowledge is not final but is in the process of development; but on the other hand, half of the students claimed that scientific knowledge is not tentative. The "man-made" characteristic of scientific knowledge was grossly misunderstood by most students. Very few agreed that "classification schemes are made up by scientists who organize observations of nature to fit these classification schemes -- that is, a classification scheme does not represent an inborn quality of the things being classified." Scientific knowledge was also misconstrued to be "absolutely true," and most students did not realize that facts are the beginning and end points of science. In addition, rock collecting was viewed as an examplary scientific investigation. However, students realistically understood that scientists do have differences of opinion and that theories we believe today were at one time vigorously opposed by emminent scientists.

Tactics of Science

The notion of "The Scientific Method" is probably one of the most widespread and misleading pieces of science education dogma ever conceived. Less than 10% of Saskatchewan chemistry students are aware of the fact that "scientific method is a myth which is usually read into the story after it has been completed." In the minds of adolescents, the discovery of "The Scientific Method" has given scientists a definite approved procedure which, if followed consistently step by step (usually five or seven), will solve almost any problem or answer almost any question in the realm of nature. Not inconsistent with these gross misunderstandings are the correct views held by over 85% of the students: (1) that failure can occur even if one rigidly follows "The Scientific Method," and (2) that there are many methods of solving scientific problems (presumably within the confines of "The Scientific Method").

Students fared much better when dealing with scientific experimentation. For most students, an experiment: (1) "is a set of conditions under which observations are made, (2) is used to test hypotheses,

(3) should be repeated, and (4) can validly yield results inconsistent with previous answers. At the same time however, a more sophisticated description of an experiment was rejected by over half of the respondents, "to set up a situation in which the control of variables is greater than it is in ordinary course of events." Painfully, the majority of students believed that "experimentation is principally concerned with proving the laws of nature." (Does this reflect the manner in which high school labs are carried out?)

Many students understood that: (1) a search of the literature, (2) the process of grouping observations, and (3) serendipity; all have a role in scientific investigations. On the other hand, a large proportion of students discountedintuition or creative imagination as leading to many important scientific ideas. Instead, students strongly supported the plodding systematic process of logical thought. In addition, they believed that large sums of money pumped into scientific research is the most important requirement for the progress of science.

Values of Science

There are many values explicitly or implicitly held by research scientists (accuracy, parsimony, open-mindedness, skepticism, honesty, to name just a few). Most of the respondents agreed that scientists should share their findings and discoveries, that this allows for independent confirmation, and that this is essential to scientific progress. However, the majority of students incorrectly felt that a scientist's reluctance to do so was based on the danger of exposing secret information. Many students properly expressed reasonable skepticism of published results but thought that "The majority of newly suggested theories are accepted by the scientific community." Most students affirmed the value that scientific conclusions should be based on fact, not opinion.

Some values (objectivity, practicality, simplicity) accepted by the working scientist were not understood by the high school students. They did not realize that objectivity is an unobtainable ideal towards which one constantly strives. Instead, the students erroneously believed that objectivity was a fact of scientific methodology. Moreover, these adolescents felt that a scientist is likely to be unbiased and objective in areas outside of science! A majority of students also held the misconception that research scientists believe their discoveries should have practical applications (a position consistent with the confusion over the aims of science and technology). Students do not view discoveries as solving intellectual puzzles or reducing a scientist's curiosity.

Few respondents realized that scientists use a simplicity value in their preference for an explanation or hypothesis. Instead, scientific knowledge was thought to be judged highly on a criterion of complexity -- the more complex, the better. (Parsimony is the value underlying the aim to formulate a simple set of expressions that explain diverse phenomena.)

Interaction of Science and Society

Most students knew that advances in science and technology have created, and will create, many changes in our culture, specifically in social and economic spheres. A large proportion of respondents believed that there must be a channel of communication between the scientific community and the lay public in order that the latter may make wise decisions concerning the former. Keeping the layman informed of scientific findings at his level of understanding was seen as crucial.

However, science was considered: (1) to have little value in helping solve today's social problems, and (2) to have no effect on society other

TABLE 15

NUMBER OF "I DO NOT UNDERSTAND" RESPONSES TO ITEMS IN THREE TESTS [a]

Test	Total number of items	# of items where "I do not understand" response is greater than 20%	# of items where "I do not understand" response is greater than 10% [b]
NOSS	28	2	1
SPI	135	7	20
TSAS	52	1	6

[a]Combined sample of all respondents.

[b]This includes the numbers in the previous column.

TABLE 16

"I DO NOT UNDERSTAND" RESPONSES TO ITEMS IN THREE TESTS [a]

Test	Items where "I do not understand" response is greater than 20%	Items where "I do not understand" response is greater than 10% [b]
NOSS	2, 9	27
SPI	13, 57, 61, 62, 89, 104, 120	8, 12, 14, 43, 50, 70, 76, 79 101, 105, 111, 118, 125
TSAS	37	4, 14, 20, 27, 44

[a]Combined sample of all respondents.

[b]This does not repeat the items in the previous column.

than through the technological implementation of scientific knowledge. Surprisingly enough, a majority of respondents disagreed that a democratic public "should ultimately control the support of science and the use of its achievements."

Human Needs in Science

Many students felt, as do many scientists, that "scientists should be concerned with the potential harm that might result from their discoveries." However, a large proportion of students were unaware that "winning the esteem of associates is one of the main incentives of the scientist."

An Analysis of Student "I Do Not Understand" Responses

Table 15 reflects the degree to which students did not understand the items on the three tests. Between 10% and 15% of the test items were not understood by at least 10% of the Saskatchewan chemistry students, while between 2% and 7% of the test items were not understood by at least 20% of the students. These items are identified in Table 16.

CONCLUSIONS

Quantitative analysis (*t* tests for matched pairs) indicated that both chemistry courses had a negligible effect on students' knowledge about the scientific enterprise. Except for the grade XI students studying Modern Chemistry (who accomplished a statistically significant gain in their awareness of some processes of science), chemistry in grades XI and XII had little impact on Saskatchewan students' impressions about science and scientists. A further analysis of student achievement revealed a slight advantage held by a course taught in a full year over the same course taught in a single semester.

The qualitative data supported this result but they also revealed far more information. In both chemistry programs, there was a balance between a gain in understanding and an acquisition of misconceptions. The meagre increase in understanding for the CHEM Study group tended to concern some social aspects of science, while the Modern Chemistry group slightly increased their understanding of some processes of science. Misunderstandings arose concerning, in part: (1) the role of imagination and intuition in doing science, (2) the scope of a scientist's investigation into natural phenomena, (3) the aims of science, (4) the nature of quantitative data, (5) the use of the simplicity value in scientific thinking, (6) the role of serendipity, (7) the modification of theories in light of new evidence, and (8) the meaning of induction and deduction.

Also identified were specific ideas which indicated strengths and weaknesses in students' understanding about science and scientists. Some of these ideas are described here. Over half of the students erroneously thought that: (1) scientific classification schemes are inherent in nature, (2) scientific laws are permanent truths, (3) the point of experiments is to prove the laws of nature, (4) there is such a thing as "The Scientific Method," and (5) money automatically buys scientific results. Over half of the students did not know that: (1) there are limitations to the use of scientific inquiry, (2) scientific models are similes or metaphors and do not duplicate reality, and (3) scientific knowledge is tentative. Over half of the students did not understand the meaning of induction, deduction, hypothesis, and theory. Over half of the students also confused the aims of science with the aims of technology. These results were similar to those of recent American surveys (Korth, 1969, using 1500 Grade XII students in the northwest sector of the United States; and Aikenhead, 1973c, using 850 students of a random sample of American and Canadian physics teachers). This recent investigation indicates that Saskatchewan students are not unique in their misconceptions about science and scientists.

Apparently the distorted ideas which students hold about the scientific enterprise are seldom challenged in chemistry classes. Students' superficial viewpoints also escape analysis.

IMPLICATIONS

The NOSS form SASK was found to have a low reliability. Several items in the NOSS are also found in the SPI or TSAS. For these reasons, it is recommended that the NOSS form SASK not be used in investigations using high school students.

The inclusion of the response "I do not understand" reduces the ambiguity in student responses because there tends to be less guessing. Further investigation is necessary, however, to study the effect of different response formats on reliability measures and statistical comparisons among group mean scores. A different scoring system (for example, 2 for a correct response, 1 for a "I do not understand" response, and 0 for an incorrect response) may also have a noticeable effect on reliability and statistical comparisons.

The Saskatchewan Chemistry Curriculum Committee are far more informed by the qualitative feedback than they are by the quantitative data. There was little apparent change in the average test scores of the CHEM Study and Modern Chemistry groups. However, with the use of qualitative data, one can pin point the ideas that were being properly learned as well as the misconceptions that were being acquired. Consequently, one can understand in what ways (by what knowledge) the two chemistry groups differed. The qualitative data give specific feedback which can be directly related to: (1) the objectives of the chemistry courses, and (2) needed improvements in the courses.

This study represents a different circumstance from the study reported at last year's NARST annual conference (Aikenhead, 1973a). There, the Project Physics group showed a statistically significant achievement over a non-Project Physics group. Therefore in the previous study, the qualitative data clarified the quantitative difference between the two groups, whereas in the present investigation, the qualitative data revealed differences where the quantitative analysis found none.

Science educators need a list of specific misconceptions generally held by high school students, for only then can curricula be developed or modified to overcome present deficits in student understanding. The data from this investigation greatly contribute to the validity of such a list.

Qualitative data appear to enhance the quality of the evaluation process. This improvement comes from asking such qualitative questions as: What ideas have students learned? What misconceptions did they acquire? What misunderstandings have they still retained?

APPENDIX A

	Number of Students	
	Sample	Poplation
Grade XI		
CHEM Study	238	8,171
Modern Chemistry	129	2,717
Grade XII		
CHEM Study	301	6,429
Modern Chemistry	73	1,866

	% of Students	
	Sample	Population[a]
Rural schools	56%	52%
Urban schools	16%	35%
Comprehensive schools	28%	13%

[a]figures are for all students in grades X to XII

APPENDIX B

ITEMS FOR WHICH THE AVERAGE STUDENT CORRECT RESPONSES WAS ABOVE 85% AND BELOW 50%

topic	above 85%*		below 50%		
	SPI	TSAS	NOSS	SPI	TSAS
Aims of science	46		4, 10 12, 22	51, 129	9, 22 28, 34 37
Epistemology of science:					
definitions			9, 11	62, 89	
assumptions	67			27, 44 75, 119	
hypotheses	80			40	
theories	25			50	
models	32, 34		16	31, 35	
laws	23, 24		5	14, 111	
general aspects	3, 28 102, 106		2, 21	41, 56 61, 120	2
Tactics of science:					
"The Scientific Method"	82, 124 135		3, 17 20, 25 28	107	3
tactics & strategies	26, 39 48, 83 122, 133 134		1, 13 15	30, 57	6
Values	3, 6 55, 60 102	15, 36	6, 7 8, 27	9, 64 86, 117	14, 32 42, 51
Interaction of science and society		1, 10 26, 40 41			19, 47 48
Human needs in science		5			11

*No NOSS item met this criterion.

REFERENCES

1. Aikenhead, Glen S. "The Measurement of Knowledge About Science and Scientists: An Investigation into the Development of Instruments for Formative Evaluation." Unpublished Doctoral Thesis, Harvard University, 1972.

2. ____________. "The Interpretation of Student Performance on Evaluative Tests." Paper presented at the 46 annual meeting of the National Association for Research in Science Teaching, Detroit, Michigan, March 28, 1973. (a)

3. ____________. "High School Physics Students' Misconceptions About Science and Scientists." (Mimeographed by the author, 1973). (c)

4. ____________. "The Measurement of High School Student's Knowledge About Science and Scientists." Science Education, 51 (#4, 1973), 539 - 549. (b)

5. ____________. "Course Evaluation II: Interpretation of Student Performance on Evaluative Tests." Journal of Research in Science Teaching, 11 (#1, 1974), in press.

6. Cooley, William W., and Klopfer, Leo. Test on Understanding Science. Princeton, New Jersey: Educational Testing Service, 1961.

7. Ferguson, George A. Statistical Analysis in Psychology and Education. 2nd ed. New York: McGraw-Hill, 1966.

8. Gallagher, James J. "Test Every Senior Project: Design and Preliminary Analysis." Paper presented at the 42nd meeting of the National Association for Research in Science Teaching, Pasadena, California, February 8, 1969.

9. Kimball, Merritt E. "Opinions Scientists and Science Teachers about Science." Unpublished Doctoral Thesis, Stanford University, 1965.

10. Korth, Willard W. "The Use of the History of Science to Promote Student Understanding of the Social Aspects of Science." Unpublished Doctoral Thesis, Stanford University, 1968.

11. ____________. "Test Every Senior Project: Understanding the Social Aspects of Science." Paper presented at the 42nd meeting of the National Association for Research in Science Teaching, Pasadena, California, February 8, 1969.

12. McNemar, Quinn. Psychological Statistics. 4th ed. New York: John Wiley and Sons, 1969.

13. Walberg, Herbert J. and Welch, Wayne W. "A New Use of Randomization in Experimental Curriculum." School Review (Winter, 1967).

14. Welch, Wayne W. Welch Science Process Inventory, form D. Cambridge, Mass.: Harvard Project Physics, 1966.

15. ______________. "Curriculum Evaluation." Review of Education Research, 39 (October, 1969), 429-443.