ABSTRACT
              Reported is a study to investigate problems
handicapped readers had with the Intermediate Science Curriculum
Study (ISCS) test and of a method for adapting the test so that it
might be more understandable. Seven questions related to the
1968-1969 ISCS test Probing the Natural World, Volume I were studied.
The investigations of these questions was undertaken in two stages:
(1) a pilot study using six classes in one junior high school, and
(2) the major study which involved the testing of 28 seventh-grade
ISCS classes at five junior high schools. The pilot study findings
indicated the need for a revised test. In addition, achievement over
the school year was examined using a repeated measure design. Item
analysis data were used to see what changes in the psychometrics of
the two instruments occurred as a result of revision. The Hotelling
t-test was used to see if the revised ISCS test had a reduced
correlation with reading as compared with the 1968-1969 ISCS test.
Reading comprehension was found to be an important factor related to
achievement on both the original and revised tests. The findings
indicated that the oral-demonstration technique is a valid method to
compensate for lack of reading ability. (Author/EB)

READING COMPREHENSION AND THE MEASURE OF SCIENCE

ACHIEVEMENT USING THE 1968 - 1969 ISCS TEST

AND THE REVISED ISCS TEST

Part III of

An Analysis of Selected Aspects of the

ISCS Model of Science Teaching

by

David William Allen

School District of Philadelphia

Introduction

This paper is an attempt to show how a written science test may be adapted so that it is more understandable for handicapped readers.

Problems

This study involved the following problems:

1. Is the 1968-1969 ISCS test <u>Probing the Natural World</u>, Volume I, an appropriate instrument for measuring science achievement in the case of seventh grade pupils two or more grade levels behind in reading comprehension in the Philadelphia Public Schools?

2. Does oral reading of the 1968-1969 ISCS test by the test administrator help handicapped readers overcome their difficulty in understanding the test items?

3. Does redesigning the 1968-1969 ISCS test produce an instrument which indicates that an oral-demonstration presentation helps handicapped readers to understand those items better than ones they read silently?

4. Does redesigning the 1968-1969 ISCS test produce
an instrument which is less dependent on reading
comprehension as measured by the 1969 Iowa Tests
of Basic Skills?

5. How do test performances on the original ISCS
test and the Revised ISCS Test compare for the
handicapped (H) and nonhandicapped (NH) readers
taking the ISCS course in the Philadelphia Pub-
lic Schools during the 1969-1970 school year?

6. Were the handicapped readers (H) and nonhandi-
capped (NH) readers' means for the experimental
oral-demonstration subtest greater than the means
for the control silent subtest of the Revised
ISCS Test?

7. Does redesigning the 1968-1969 ISCS test produce
an instrument which appropriately measures science
achievement in the ISCS course for seventh grade
Philadelphia Public School pupils?

Procedures

The investigations of these seven questions was under-
taken in two stages: (1) a pilot study using six classes in
one junior high school, and (2) the major study which in-
volved the testing of 28 seventh grade ISCS classes at five
junior high schools. Teachers whose classes were used in

the study participated in the ISCS in-service workshop at
Temple University and volunteered in March 1970 to have their
classes tested in June using the Revised ISCS Test. Test
data on the same sample using the original ISCS test was made
available from a companion study by Clark (1973) to answer
questions 1, 4, and 5.

The pilot study investigated problems 1, 2, and 3 to
find out if a problem with readability of the original ISCS
test was present. A junior high school with two ISCS teachers,
each having three seventh grade ISCS classes, was used for the
pilot study. In order to answer question 1 mean reading com-
prehension levels were computed for each class. Reading com-
prehension scores were obtained from the 1969 administration
of the Iowa Tests of Basic Skills (ITBS). Classes which had
mean reading comprehension two or more years below the readabil-
ity level of the original ISCS test of seventh grade third
month as calculated using the Dale-Chall Formula were categor-
ized as having difficulty reading the original ISCS test.

For classes which were categorized as handicapped in
their ability to read the original ISCS test, a t-test was
calculated to determine if H classes did better on the March
ISCS posttest than the chance scores of 11 for a 44 item
multiple choice test. This March ISCS posttest was admin-
istered by the teachers as part of Clark's companion
study.

In order to determine whether the mere oral reading
of the original ISCS test might help handicapped readers

the following design was implemented to answer question 2.

The pupils in the six ISCS classes were matched in one of two ways: (1) on self using a split-halves treatment of the original ISCS test, or (2) with a pupil of the same sex with similar academic achievement. The experimental treatment consisted of an oral presentation of some of the items on the 1968-1969 version of the ISCS test. In the control presentation pupils read items silently as they had in the March ISCS posttest. It was hypothesized that the oral reading might lead to higher scores on the oral subtest as compared to the silent subtest, particularly for H reading classes. Nonparametric tests were used to test for the significance of the difference between the experimental and control groups. Failure to find significant differences favoring the oral treatment for H classes would be interpreted as evidence that the mere oral reading of the original ISCS test was not sufficient to help increase pupils' understanding of the items.

A revision of the 1968-1969 ISCS test was undertaken to make some of the items more understandable. Additional pictures were added and vocabulary and sentences were changed in an attempt to increase the readability. Half of the 38 items selected from the original ISCS test were presented orally with demonstrations using the ISCS science equipment available in the classroom. A t-test of the difference was calculated to find out if pupils performed better on the orally-demonstrated subtest as compared to the silent

subtest. If the oral-demonstration subtest mean was found to be significantly above the silent mean, the major study would investigate questions 1, 4, 5, 6, and 7.

In the major study pupils from five junior high schools were administered the Revised ISCS Test in June 1970. Pupils present for the September, March, and June administrations of the ISCS tests who had 1969 ITBS reading comprehension subscores were used in the analysis of the data. There were 188 handicapped readers (H group) and 242 nonhandicapped readers (NH group) whose data were used to answer the questions.

Class means for the 1970 ITBS reading comprehension subtest were computed to categorize classes on their ability to read the original ISCS test. If the reading comprehension grade equivalent mean score for a class was two years or more below the readability level of the original ISCS test (i.e., seventh grade third month), then that class was categorized as handicapped in reading the ISCS test and was placed in the H group. All classes above the cut-off point of fifth grade third month were considered to be nonhandicapped in their ability to read the original ISCS test. These classes are referred to as the NH group.

To answer question 1 to see if the conclusion in the main study replicated the pilot study findings the fiducial limits of the original ISCS test were computed for the H group. The investigator wanted to know whether the H group means on the ISCS test were significantly different from

the chance means of the September and March administrations of the original ISCS test. Failure to find a positive significant difference in the March posttest would be considered as evidence of the inappropriateness of the 1968-1969 version of the ISCS test for handicapped readers.

Correlations were computed between the two ISCS tests and the 1969 reading comprehension ITBS score to answer question 4. Since both the original ISCS test and the Revised ISCS Test required some reading it was of interest to see if the correlations with reading were different for the two forms of the test. An analysis of H group data was performed to see if a difference between the March original ISCS test correlation with reading and the Revised ISCS Test correlation with reading was significant using a Hotelling t-test. In addition a comparative study of H and NH groups was undertaken to see if science achievement differed significantly between the different reading ability groups.

Comparisons of the H and NH groups' achievement as measured by the original and the Revised ISCS tests were made using analysis of variance to compute an F ratio. A significant difference in achievement favoring the NH group (i.e., able reading classes) would be interpreted as evidence that reading comprehension ability was an important factor which may have limited the amount of science learning that was measurable using the ISCS tests. The results from the analysis of H and NH groups' science achievement would be used to answer question 5.

To answer question 6 the Revised ISCS Test was designed with two subtests: (1) an oral-demonstration subtest, and (2) a silent subtest. The oral-demonstration items were alternated with silent items to determine if there was a treatment effect resulting from the oral-demonstration presentation. The oral set and silent set of items, which were taken from the 1968-1969 form of the ISCS test, were of like difficulty according to March posttest analysis (i.e., D = .41) for the NH group, but had a slight difference in difficulties of .05 for the H group, with the oral set being slightly easier. Thus in the case of the H group there may have been a slight difference in difficulties between the two subtests before the items were revised and the experimental oral-demonstration administration was applied. It was presumed that revisions of both the oral and silent items would lead to changes in difficulty. However, in addition to the written revisions, the oral-demonstration presentation may have made certain items less difficult.

A t-test of the difference between the oral-demonstration subtest mean and the silent subtest mean was used to see if the treatment aided H and NH pupils in understanding the items better, enabling them to achieve higher scores on the oral subtest than the silent subtest. If the Revised ISCS Test means for the H and NH groups were above the chance level mean of the test and if the experimental oral treatment mean was significantly greater than the nontreatment silent subtest mean, then there was some evidence that

the oral-demonstration technique contributed to the design-
ing of an appropriate instrument.

A questionnaire was given to H and NH classes to com-
pare pupils' feelings toward the original and Revised ISCS
tests. The results of the questionnaire were utilized to
answer question 7 in conjunction with H group test results.

In addition a repeated measures design was used to
see if the H group made significant gains in science in
order to answer question 7. September, March, and June ad-
ministrations of ISCS tests were used to measure science
achievement during the 1969-1970 school year. If the over-
all F ratio was found to be significant, then the mean dif-
ferences between September and March, and March and June
tests would be analyzed using a Tukey test. Failure to show
significant gains in H group pre-to-posttest means would be
interpreted as due to the inappropriateness of one or both
instruments.

Pilot Study Findings

Question 1. Is the 1968-1969 ISCS test Probing the
Natural World, Volume I, an appropriate instrument for
measuring science achievement in the case of seventh grade
pupils two or more grade levels behind in reading compre-
hension in the Philadelphia Public Schools?

No, the 1968-1969 version of the ISCS test was in-
appropriate for measuring the science learned by H classes.
The March ISCS posttest means for the two handicapped

classes were found to be within the chance distribution of test scores (i.e., see Table 3 in Appendix).

Question 2. Does oral reading of the 1968-1969 ISCS test by the test administrator help handicapped readers overcome their difficulty in understanding the test items?

No, handicapped reading classes were not aided by orally presenting parts of the original ISCS test. As indicated on Table 4 in the Appendix, results from nonparametric tests of the difference between oral and silent treatments revealed that the H classes were not helped by the test administrators' oral presentation of certain items. One H class did significantly better on the silent subtest.

Question 3. Does redesigning the 1968-1969 ISCS test produce an instrument which indicates that an oral-demonstration presentation helps handicapped readers to understand those items better than ones they read silently?

Yes, the oral-demonstration subtest mean scores were significantly above the silent subtest mean for the H classes as indicated on Table 6 in the Appendix.

Major Study Findings

Question 1 was asked again to see if the finding of the pilot study was replicated in the major study.

Question 1. Is the 1968-1969 ISCS test Probing the Natural World, Volume I, an appropriate instrument for

measuring science achievement in the case of pupils two or
more grade levels behind in reading comprehension in the
Philadelphia Public Schools?

No, the 1968-1969 version of the ISCS test was in-
appropriate for measuring the science learned by the H
group. This finding was in agreement with the pilot study
conclusion. As indicated on Table 13 in the Appendix, the
ISCS mean for the H group on the March posttest using the
original ISCS test was not statistically significant in its
difference from the chance score of 11 for the 44 item test.

A gain score analysis of September and March ISCS
test results are presented in Table 16. A Tukey test of the
difference between March 1970 and September 1969 means using
the original ISCS test was not significant. This further
supports the conclusion that the original ISCS test was not
appropriate for measuring progress of H classes taking the
ISCS course.

Question 4. Does redesigning the 1968-1969 ISCS
test produce an instrument which is less dependent on reading
comprehension as measured by the 1969 Iowa Tests of Basic
Skills?

No, the Revised ISCS Test was found to be signifi-
cantly more correlated with reading comprehension than the
original ISCS test. The results of the Hotelling t-test are
in Table 23. The correlation of the 1969 ITBS reading com-
prehension grade equivalent scores increased significantly

from the administration of the original ISCS test in March
and the Revised ISCS Test administered in June. This find-
ing was a surprise; however, it was not known when the study
was undertaken that the correlation between the 1969 ITBS
reading comprehension subtest and the 1968-1969 form of the
ISCS test was near zero. Restriction of the range of pos-
sible ITBS scores as a result of selecting classes in the
H group on the basis of their being two years below the
readability level of the ISCS test may have been part of
the reason for the near zero correlation. In addition,
random test taking behavior on either the March ISCS test
or the 1969 ITBS test might have contributed to the low cor-
relation between a science and a reading comprehension test.

Question 5. How do test performances on the original
ISCS test and the Revised ISCS Test compare for the handi-
capped (i.e., H) and nonhandicapped (i.e., NH) readers taking
the ISCS course in the Philadelphia Public Schools during
the 1969-1970 school year?

There was a significant difference between H and
NH performance on the September 1969, and the March and
June 1970 administrations of ISCS tests as indicated in
Tables 17 and 18, and Figure 2. The significant difference
in performance between the H and NH groups on three adminis-
trations using two versions of the ISCS test indicated that
reading comprehension achievement levels were an important
factor affecting pupil ability to both learn and show what

they had learned from their ISCS course.

Question 6. Were the handicapped readers (H) and nonhandicapped (NH) readers' means for the experimental oral-demonstration subtest greater than the means for the control silent subtest of the Revised ISCS Test?

Yes, there was a treatment effect favoring the oral-demonstration technique over the silent treatment as indicated in Table 20. The mean for the oral-demonstration subtest was significantly greater than the mean for the silent subtest for both H and NH groups. This occurred in spite of the fact that many of the items were paired, so that if the oral item came first, pupils may have been helped in their understanding of the next silent item. This treatment effect aided in making orally-demonstrated items more understandable and may have had a spin-off effect of making some of the related silent items more understandable.

Question 7. Does redesigning the 1968-1969 ISCS test produce an instrument which appropriately measures science achievement in the ISCS course for seventh grade Philadelphia Public School pupils?

Yes, pupil performance above the chance level of the Revised ISCS Test administered in June indicates that the redesigned test was appropriate. The fiducial limits on the test performance data for the H group subtests are presented in Table 21.

In addition a gain score analysis using the Tukey

test revealed that significant gains had been made between September and June and March and June, but not between September and March. Thus, the use of the Revised ISCS Test in June was an important factor in picking up significant gains indicative that some science learning did occur. The results of gain score analysis are presented in Table 16.

Pupils' responses to a questionnaire also revealed that they were able to understand the items better on the Revised ISCS Test as compared to the original ISCS test.

Summary

The dissertation has investigated the problem of the appropriateness of the ISCS test and adapted the 1968-1969 version of the ISCS test so that it might be better understood by handicapped readers who took the seventh grade ISCS course in the Philadelphia Public Schools during the 1969-1970 school year. Reading comprehension was found to be an important factor related to achievement on both the original and Revised ISCS tests. The findings in this dissertation indicate that the oral-demonstration technique for presenting written items is a valid method to insure that the lack of reading ability does not block the measurement of science learned. Learning to read science materials needs to be a major objective of all science teachers if written science tests are to be used as the criterion for science learning.

Some significant, although possibly low, correlation should exist between a written science test (some of which

may be presented orally) and a reading comprehension test in

order to obtain a meaningful measure of science learning.

The fact that NH classes also improved significantly

on the Revised ISCS Test was interpreted as evidence that

some of the pupils in those classes were also having diffi-

culty in reading the original ISCS test. Thus, the investi-

gator concluded that the special oral-demonstration would

be appropriate for pupils whose reading comprehension level

was one year below the readability level of the test.

The rewording of items and addition of pictorial

representations to 11 of them, plus the oral-demonstration

presentation, indicated to pupils that the test administrator

really cared about the science they had learned. Thus,

the outcome was a response pattern other than the random

filling-in of the answer sheet. More humane techniques

for administering written tests to handicapped readers

need to be utilized if valid measures of learning are

desired.

APPENDIX

TABLE 3

MARCH ISCS MEANS FOR HANDICAPPED READING CLASSES AND t TEST

Statistic	MARCH ISCS TEST		
	Class 1	Class 4	Class 1 + Class 4
Number of Students	24	22	46
Test means	13.3	11.3	12.3
Standard deviation	3.48	3.19	3.60
Chance level (44 items)	11	11	11
Difference from chance	2.4	.3	1.4
t value (d.f.)	3.01 (23)	.397 (21)	2.68 (45)
Probability	S.	N.S.	N.S.
Level of rejection[a]	.01	.01	.01

[a]The .01 level was chosen for H_{o_1} to reduce the risk of a Type I error.

TABLE 4

MARCH ISCS SUBTEST MEANS AND TEST FOR THE SIGNIFICANCE
OF THE DIFFERENCE BETWEEN ORAL AND SILENT SUBTESTS

Class Match-ing	Numbers of Items Ad-ministered	Class Means on ISCS Subtests			Test of Significance (p<.05)
		N	Oral	Silent	
1 p[a] H	23	7	8.09	9.36	-S.
2 s[b]NH	36	25	7.64	6.11	+S.
3 p NH	23	9	12.11	10.11	N.S.
4 p H	23	9	8.50	8.16	N.S.
5 s NH	36	25	7.04	6.52	N.S.
6 s NH	36	26	9.41	12.26	N.S.

[a]Matched pairs of pupils HANDICAPPED in reading

[b]Pupils received a split half treatment, and therefore are matched on self, NONHANDICAPPED in reading.

TABLE 6

READING HANDICAPPED GROUP
t TEST OF THE DIFFERENCE BETWEEN ORAL AND SILENT SUBTESTS
OF THE REVISED ISCS TEST

$N = 46$

Mean Difference = .98

Standard Deviation of the Mean = 2.62

Standard Error of the Mean = .39

t Value = 2.51

d.f. = 45

$p < .05$

TABLE 7

REVISED ISCS TEST MEANS FOR PILOT STUDY CLASSES

Statistic	1	2	3	4	5	6	Overall
				Class			
Number of Pupils	24	27	23	20	25	25	143
Test Means	17.0	17.5	17.6	11.2	15.6	21.5	16.6
Oral Mean	8.6	7.7	9.2	6.6	8.1	12.1	8.8
Silent Mean	8.4	9.8	8.4	4.6	7.5	9.4	7.8
Difference	.2	-2.1	.8	2.0	.6	2.7	1.0

TABLE 13

TEST OF FIDUCIAL LIMITS OF ISCS MEANS
FOR HANDICAPPED (H) READERS

Statistic	Administration		
	September	March	June
N	188	188	188
M	10.10	11.35	13.6
Chance Score	11	11	9.5
Z	- 4.5^a	.1.67	13.6
P	**	N.S.	**

** p < .01

^{a} May be indicative that the original ISCS test for H group was a speeded test rather than a mastery test.

TABLE 16

ANALYSIS OF HANDICAPPED READING CLASSES
GAIN SCORES OVER THREE TIME PERIODS

| Time Periods	2-1	3-1[a]	3-2
Mean Difference	1.25	3.50	2.25
Q Values - Critical $_{.95}Q = 3.53$	2.41	4.25	3.70
Level of significance	N.S.	*	*

*$p < .05$

[a]This gain score analysis is based on the assumption that the two
forms are paralled.

TABLE 17

A COMPARISON OF ISCS TEST MEANS FOR HANDICAPPED (H)
AND NONHANDICAPPED (NH) READERS

Administration	Blocking	N	Statistic M	S.D.
September	H	188	10.0	3.52
	NH	242	11.8	3.67
March	H	188	11.3	3.86
	NH	242	16.4	3.91
June	H	188	13.6	3.81
	NH	242	19.4	4.76
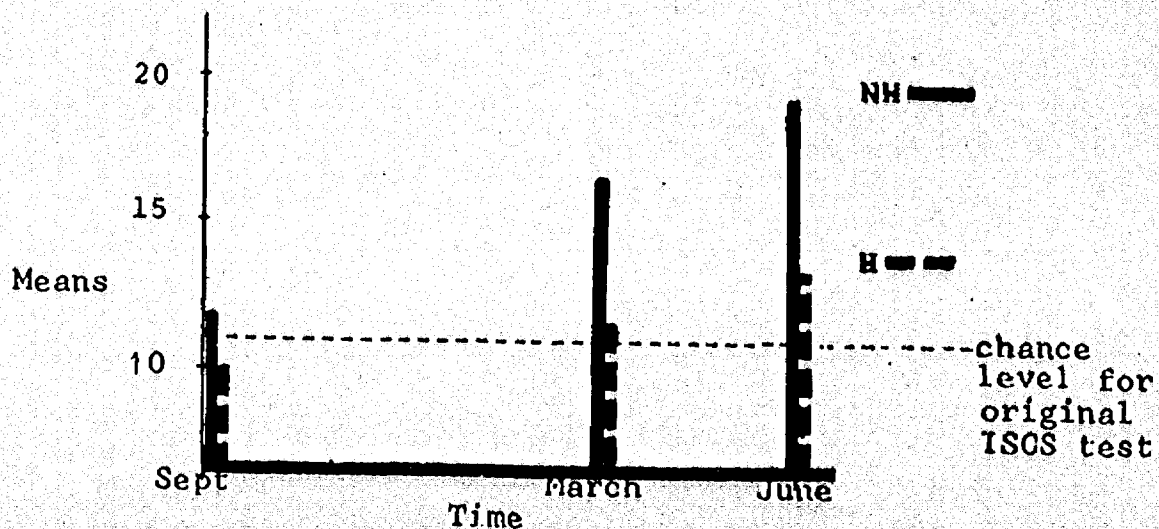
[1] A quadratic component was found to be present for the H group.
[2] Linearity of regression was tested and found to be signifi-
cant for both the NH and H groups at the .01 level.

FIGURE 2

A COMPARISON OF HANDICAPPED AND NONHANDICAPPED
READERS' ISCS ACHIEVEMENT[1,2]

TABLE 18

RESULTS OF ANALYSIS OF VARIANCE

ON PRE- POST- AND REVISED ISCS MEAN SCORES OF H AND NH GROUPS

TO TEST NULL HYPOTHESIS 4

Source of Variation	SS	df	MS	F	P
Within Cells	72.68	26	2.79		
Between Groupings	165.54	1	165.54	59.2	.001

TABLE 20

TEST OF SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS ON
TWO SUBTESTS OF THE REVISED ISCS TESTS USING A
CORRELATED t TEST

Variables	Handicapped	S.D.	Non Handicapped	S.D.
Mean for Oral Test	7.33	2.57	10.11	2.79
Mean for Silent Test	6.26	2.17	9.40	2.77
Difference	1.07		.71	
Number of Pupils	188		242	
Pearson r	.88		.45	
t value	4.3**		3.6**	

** p < .01

TABLE 21

TEST OF FIDUCIAL LIMITS OF THE REVISED ISCS SUBTESTS
FOR THE H GROUP

	Subtest	
Statistic	Oral	Silent
Number of Pupils	188	188
Mean for the Subtest	7.33	6.26
Chance Score	4.50	4.50
Z Value	14.90	11.00
Probability Level	**	**

** p < .01

TABLE 23

TEST OF SIGNIFICANCE BETWEEN TWO CORRELATION COEFFICIENTS

$.12_r$ ISCS, Iowa Reading $-.29_r$ Revised ISCS, IOWA Reading	$= -.17$
r ISCS, Revised ISCS	$= .28$
Number of pupils taking all three ISCS tests	$= 188$
d.f.	$= 185$
t_{dr}	$= -2.01*$

*$p < .05$