ABSTRACT
              The Kilander Health Knowledge Test is analyzed, based
on responses of four independent groups, to determine if response
tendencies to test items were peculiar to groups studied and if test
items retained independence regardless of the population group..The
1969 edition of the Kilander Health Knowledge Test was administered
to all ninth- and twelfth-grade students attending a high school
representing both a high and low socioeconomic level. The test was
deemed to be sensitive to response tendencies of independent
population groups. Analysis procedures also established that test
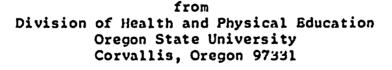item categories were independent regardless of the population.
(JB)

ITEM ANALYSIS OF KILANDER HEALTH KNOWLEDGE
TEST USING FOUR INDEPENDENT POPULATIONS 1/

by

Donald E. Campbell
and
Roy A. Foster

ITEM ANALYSIS OF KILANDER HEALTH KNOWLEDGE
TEST USING FOUR POPULATION GROUPS

In a series of provocative papers, Berg (1, 2) has insisted

that individuals taking a test make many responses which are simi-

lar to responses made by others, but they also make responses

which are unlike those of others in the general population. Res-

ponses in the second category were reported to be most like those

made by members of some special sub-population to which that parti-

cular individual belonged. Berg concluded that test content is

secondary to clearly delineated criterion groups and an analysis

of response tendencies peculiar to the particular groups.

Investigators have reported the use of the Kilander Health

Knowledge Test as the criterion instrument to assess health know-

ledge of various population groups. In spite of the fact that

significant difference in group means has been reported, little

evidence has been presented to establish that the difference in

mean scores represent a difference in knowledge level and not a

difference in response tendencies peculiar to sub-populations

studied.

The purpose of this investigation was to conduct an item

analysis of the Kilander Health Knowledge Test based on the res-

ponses of four independent population groups in order to determine

if response tendencies to test items are peculiar to groups studies

and if test items retain item independence regardless of the

population group.

METHODOLOGY

Criterion Instrument and Measures

The 1969 edition of the Kilander Health Knowledge Test was used as the criterion instrument and the criterion measures consisted of the total correct response score, nine category scores, and item response to each of the 100 questions.

This test was designed for high school and college students and consists of 100 multiple choice questions. Questions were selected from nine aspects of health education; however, a few items test attitudes rather than knowledge. Reliability of .80 for college freshmen and .83 for high school seniors has been reported. Scores ranging from 22 to 82 and a mean of 56.15 have been reported for high school seniors. Several investigators have used the instrument with high school populations (5, 6, 7, 10, 12); however, the instrument has more often been used with college populations (3, 4, 8, 12, 13, 14, 15). Test item discrimination has not been reported.

Criterion Groups

During the fall of 1971 the criterion instrument was administered to all of the ninth and twelfth grade students attending two senior high schools in the City of Portland school district. The designation of the two schools was made by the central administration and approved by the principal of the school involved. The criterion for the selection of participating schools specified

that the student body of a school would represent a high socio-
economic level or a low socioeconomic level but would otherwise
reflect similar curriculum opportunities, an equal balance in
racial and ethnic population, a comparable student-teacher ratio,
and proportional size. Socioeconomic level was to be determined
by estimated average income, occupation of parents, and value of
residence (9, 16). The two selected schools represented a reason-
ably clear delineation of the criterion. The literature suggested
that difference in response might be expected by grade level as
well as socioeconomic level (5, 6). Thus four criterion groups
were established to include a low socioeconomic level ninth grade
(LSL-9), low socioeconomic level twelfth grade (LSL-12), high socio-
economic level ninth grade (HSL-9), and high socioeconomic level
twelfth grade (HSL-12). A significant F value between socio-
economic level (F=139.56) and between grade levels (F=195.12) has
been reported as has a significant interaction effect (F=5.74) (6).
Therefore, the four groups were regarded as four independent popula-
tion groups.

Test Administration and Scoring

The standard procedures for the administration of the Kilander
test were precisely executed, except that each student recorded the
response to each of the 100 items on an IBM 503 answer sheet instead
of marking the test booklet. The test was administered to the
ninth grade students by the ninth grade English teachers and to the
twelfth grade students by the twelfth grade social science teachers.

Each classroom teacher was assisted by a member of the research team. The investigators met with each classroom teacher prior to the administration of the test to standardize testing procedures and to distribute test materials (6).

The completed answer sheets were processed by an IBM 1232 Optical Mark Page Reader and a 534 Card Punch to provide punched data tab cards for each subject's responses. These cards were subsequently submitted to a CDC 3300 computer with a program to obtain punched data cards containing the total and the category scores. These data cards were used to obtain distribution statistics, item analysis data to include Difficulty Index and Discriminant Index, distribution of responses to include X^2 values for independence of distribution, and correlation coefficients of the criterion scores.

RESULTS AND DISCUSSION

The mean score and distribution statistics for the responses on the Kilander Health Knowledge Test by four independent population groups are presented in Table 1. Scores for each population group were arranged in order of magnitude. Respondents with scores in the upper and lower 27 percent of the range were selected from computation of the Discriminant Index and Difficulty Index (11). The number of respondents in the lower 27 percent of the range of scores represented an equal proportion of each criterion group. The number in the upper 27 percent of the range of scores also represented an equal proportion with the exception of the LSL-9 group which had a smaller proportion of the total group in the

upper 27 percent. A comparison of the range of scores, median, mean, and standard deviation would suggest that the socres for the LSL-9 group were skewed to the low side.

Distribution of responses, percent of correct responses for the total group, and for both the upper and lower sub-groups, Difficulty Index, and Discriminant Index were obtained for each of the 100 items for each of the four independent groups. The results of item number 9, which had the highest Difficulty Index, and item number 91, which had the lowest Difficulty Index are presented in Table 2. Both items had a significant $\underline{x}^2$ value for independence of distribution of responses. Neither item had a significant $\underline{x}^2$ value for independence when the two twelfth grade groups were considered separately.

The difference between the means for the LSL-9 and LSL-12 was 12.89, while th. difference between HSL-9 and HSL-12 was 9.38. This would sugeest that the students in the LSL school gained more information between the ninth and twelfth grades than did the students in the HSL school. Students in the LSL-12 scored only slightly higher than did the HSL-9 students.

Test item 9 dealt with the specialization of physicians. The students in the LSL school had only a change of 2 points in difficulty between the ninth and twelfth grades while the students in the HSL school had a change of 17 points between the ninth and twelfth grades. This would indicate that the students in the HSL group gained more knowledge about specialization of physicians

than did the students in the LSL group. Test item 91 dealt with

mental disorders. The low Difficulty Index would indicate that

this was a difficult question for all students.

The ten least difficult and the ten most difficult items on the

Kilander Health Knowledge Test as determined by the Difficulty Index

are reported in Table 3. Difficulty Index was determined on the

bases of the percent of correct answers by respondents in the upper

27 percent of the range as contrasted to percent of correct answers

by respondents in the lower 27 percent as both were contrasted to

the percent of correct replies by the total group. A high index

score established items which were answered correctly by most res-

pondents while a low index score indicated a tendency to have an

incorrect response.

Items 9, 29 and 55 (Table 3) were included by all four criter-

ion groups when the ten items with the highest Difficulty Index

were selected for each of the four groups. These items represented

the least difficult items. The remaining seven items in Table 3

represented items which had the highest averaged Difficulty Index

score across four criterion groups. Some of these items were not

included in the ten least difficult items for the specific group.

Items 91, 63, 97 and 88 in the Low Difficulty Index column of

Table 3 were items to have a common low Difficulty Index when the

ten items with the lowest difficulty index were selected for each

criterion group. These four items represented the most difficult

items for all respondents. The remaining six items represented

questions which had the lowest averaged Difficulty Index score

across the four groups.

In addition to the ten items shown in Table 3, items 14, 71, 80 and 85 were particularly difficult for the LSL-9 group, while items 4, 5, 94 and 95 were additional difficult questions for the HSL-9 group. The LSL-12 group seemed to have difficulty with items 4, 11 and 83 while the HSL-12 group had items 66 and 83 in the ten most difficult list.

Items with considerable variation in Difficulty Index by the four population groups are presented in Table 4. The distribution of responses by population groups, x^2 for independence of distribution of responses, and correct response are also presented. Responses to these ten items seemed to account for much of the difference in groups means.

Analyses of the test items found in both the high Difficulty Index and low Difficulty Index indicated that no consistency in health topics existed in either group. The difficulty of the test item appeared to be within the item rather than within the health topic. Chi square values for these items established that all but item 45 and 63 had independence of distribution of responses. Test items 50, 69, 73, 83 and 94 had an increase in difficulty from the ninth to the twelfth grade. Since each of these items pertained to different health topics, the increase in difficulty would seem to be due to the item rather than to a deficiency in knowledge related to specific health topics.

Four items with a common high Discriminant Index and four items with a common low Discriminant Index are presented in Table 5.

Discriminant Index is the difference between the percent of correct answers for respondents in the upper 27 percent of the range of scores and the percent of correct answers for the respondents in the lower 27 percent of the range. The Index represented the difference in the extent to which the upper 27 percent consistently marked the item correct and the lower 27 percent consistently marked the item correct. A high Discriminant Index would indicate that the upper group had a high percent of correct response and the lower group had a low percent of correct response. A low Discriminant Index would indicate that the difference between the extent to which the upper group gave the correct response and the lower group gave the correct response was minimal. A negative index score would indicate that the respondents in the lower 27 percent of the range of scores gave the correct response more consistently than those in the upper 27 percent of the range of scores.

Items 35, 49, 54 and 61 had high Discriminant Index for ninth graders but not twelfth graders, while items 14, 48, 80 and 89 had high Discriminant Index for twelfth graders but not ninth graders. The LSL groups had high Discriminant Index on item 42 and the HSL groups had a high Discriminant Index on item 100. Item 95 had a common low Discriminant Index for both ninth grade groups but not twelfth grade groups. In addition, to the four items shown in Table 5, the LSL-9 group had 16 items with low Discriminant Index, while LSL-12 had only 4, HSL-9 only 5 and HSL-12 only 2 other items.

Items 94 and 96 had a common negative Discriminant Index across

all groups. The LSL-9 group had six additional items (45, 53,

65, 81, 90 and 93) with a negative Discriminant Index. The HSL-12

group had a negative Discriminant Index on items 68 an. 93.

The Discriminant Index indicated very little difference bet-

ween the two high schools. The ninth grade group in the LSL

school had more items with a low Discriminant Index and more with

a negative index than any of the other groups. This would indicate

that for those items the students in the ninth grade of the LSL

high school had less knowledge than the other students. The

twelfth grade students in the same school had considerably fewer

items with a low or negative Discriminant Index. These results

suggest that between the ninth and twelfth grades the students

in the LSL high school gained more health knowledge than did the

students in the HSL high school.

Chi square analysis of the distribution of the four possible

responses for each of the four groups for each of the 100 test items

was computed from a 4 by 4 contingency table in order to determine

if an equal number of respondents would have the same distribution.

Independence values were first computed for the expected frequences.

These independence values were obtained by multiplying the fre-

quences of the rows and columns for a given cell and dividing this

value by N. The chi square for each cell was then computed in the

usual manner. Chi square necessary for significance with 9 df at

.01 level was 21.66. Items 5, 8, 12, 15, 16, 23, 27, 45, 55 and 63

were the only items for which a significant x^2 value for independence of distribution was not obtained. Five of these items were from the Personal Health category of the test and the others were distributed across the remaining categories.

Correlation of category scores to total score was obtained for each of the four population groups for each of the nine categories of the test (Table 6). While the significance of the difference between correlation coefficient values for the four independent samples was not obtained, no consistent pattern of difference or r for category to total scores was apparent. Regression coefficient of the category to total score was also computed with the results shown in Table 6. Perusal of the regression coefficient across criterion groups will suggest common values for each dependent measurement.

While the Discriminant Index, Difficulty Index and x^2 values for independence of response would suggest that the four populations did respond to items with some degree of independence, the consistency of the correlation coefficient and regression coefficient values suggested that the test items had common response tendencies within the population groups.

Using a full selection of response scores, a biserial coefficient correlation item analysis procedure was used to compare the distribution of item scores to the total score as well as the item score to the category score for all of the items in the Family Living, Mental Health, and Safety Education categories.

On the basis of a multivariate discriminate analysis, response scores on these three categories had been reported to be the primary contributors to the separation of the four population groups (6). While no consistent pattern of difference was apparent, a considerable difference in r_{bis} values will be observed.

Within each category an item-to-item correlation coefficient was also obtained. For the nine Family Living category items, a low item-to-item correlation was found for all items for the four criterion groups with the exception of items 61 to 64, 61 to 74, and 64 to 74 for the LSL-9 group. For the five Mental Health items, a low r for all items was obtained with the exception of items 50 to 56 for both the ninth and twelfth grade LSL groups. Four the three Safety Education items a low r for all items across four independent groups was obtained. These results would suggest that with the possible exceptions cited that the test items within the category were independent.

SUMMARY AND CONCLUSION

On the basis of the item analysis procedures and population groups used in this investigation, the Kilander Health Test was deemed to be sensitive to response tendencies of independent population groups and at the same time to have a high degree of consistency within the population groups. Analysis procedures also established that test items within categories were independent regardless of the population.

TABLE 1.

MEAN SCORE AND DISTRIBUTION STATISTICS FOR KILANDER HEALTH KNOWLEDGE TEST FOR FOUR GROUPS

Variable	Groups			
	LSL 9	LSL 12	HSL 9	HSL 12
N	228	254	362	322
Mean [a]	33.30	46.19	44.48	53.86
S.D.	10.01	12.89	12.44	14.00
Median	37	48	44	57
Range	20-69	18-84	11-84	10-87
Number in lower 27% of range of scores	68	76	110	96
Highest test scores of lower 27%	32	40	37	48
Number in upper 27% of range of scores	47	66	103	94
Lowest test score of upper 27%	44	56	52	65

[a] A significant F value at .01 level for Socioeconomic level and Grade level and a significant interaction effect at .05 level has been reported (6).

TABLE 2.

ANALYSIS OF KILANDER HEALTH KNOWLEDGE TEST ITEMS WITH
HIGHEST AND LOWEST DIFFICULTY INDEX BY FOUR GROUPS

Test Item	Correct Response	Distribution of Responses 1	2	3	4	NR	Group		Percent Correct	Difficulty Index	Discriminant Index
										High Difficulty Index	
							LSL-9				
		167	26	25	10	0	Total	73			
		45	11	9	3	0	Lower	66		76	19
		40	3	2	2	0	Upper	85			
							LSL-12				
		210	17	19	8	0	Total	83			
		47	11	11	7	0	Lower	62		78	32
		62	3	1	0	0	Upper	94			
9	1						HSL-9				
		268	29	43	20	2	Total	74			
		63	17	20	9	1	Lower	57		72	30
		90	1	10	2	0	Upper	87			
							HSL-12				
		291	10	18	2	1	Total	90			
		78	9	7	1	1	Lower	81		89	16
		91	0	3	0	0	Upper	97			
										Low Difficulty Index	
							LSL-9				
		148	39	28	12	0	Total	12			
		49	10	9	0	0	Lower	0		4	9
		25	13	4	4	0	Upper	4			
							LSL-12				
		117	65	32	38	2	Total	15			
		35	17	12	12	0	Lower	16		18	4
		30	19	4	13	0	Upper	20			
91	4						HSL-9				
		77	134	76	71	4	Total	20			
		23	37	23	23	4	Lower	21		21	-1
		20	45	17	21	0	Upper	20			
							HSL-12				
		172	88	34	27	1	Total	8			
		63	16	12	4	1	Lower	4		11	13
		40	30	8	16	0	Upper	17			

TABLE 3.

KILANDER HEALTH KNOWLEDGE TEST ITEMS WITH COMMON HIGH AND LOW DIFFICULTY INDEX BY FOUR SUBJECT GROUPS

High Difficulty Index					Low Difficulty Index				
Test Item	Groups				Test Item	Groups			
	LSL 9	LSL 12	HSL 9	HSL 12		LSL 9	LSL 12	HSL 9	HSL 12
9	76	78	72	89	91	4	18	21	11
29	72	77	77	79	63	17	25	16	17
55	83	74	75	82	97	13	18	30	24
58	62	75	73	84	88	12	21	25	19
1	62	78	77	85	18	21	27	18	39
26	60	75	67	81	3	14	29	21	31
35	65	73	67	83	87	13	34	33	44
90	75	72	58	85	95	12	24	28	39
25	66	67	73	75	53	26	29	23	31
30	60	62	62	75	45	27	31	28	31

TABLE 4.

DISTRIBUTION OF RESPONSES ON TEN KILANDER HEALTH KNOWLEDGE
TEST ITEMS WITH CONSIDERABLE VARIATION IN
DIFFICULTY INDEX ACROSS FOUR GROUPS

Test Item	Group	Distribution of Responses					x^2	Correct Response	Difficulty Index
		1	2	3	4	NR			
	LSL 9	60	55	70	43	0			25
4	LSL 12	64	33	71	85	1	106.76	1	30
	HSL 9	83	92	100	86	1			25
	HSL 12	167	42	53	60	0			52
	LSL 9	53	23	101	51	0			43
13	LSL 12	33	23	150	45	3	101.71	3	62
	HSL 9	53	52	160	96	1			48
	HSL 12	18	12	244	48	0			72
	LSL 9	140	42	27	19	0			25
	LSL 12	104	118	24	7	1	160.68	2	49
14	HSL 9	225	70	43	22	2			25
	HSL 12	102	187	27	6	0			61
	LSL 9	64	89	15	59	1			44
	LSL 12	40	125	19	68	2	57.37		53
38	HSL 9	67	186	30	78	1		2	54
	HSL 12	38	219	15	50	0			68
	LSL 9	80	83	41	22	2			38
	LSL 12	83	101	53	15	2	59.94		34
50	HSL 9	110	149	78	23	2		1	29
	HSL 12	180	82	48	12	0			61
	LSL 9	70	67	13	76	2			30
	LSL 12	27	156	15	54	2	99.12		57
65	HSL 9	37	194	24	106	1		2	52
	HSL 12	34	194	6	86	2			58
	LSL 9	140	43	37	8	0			66
	LSL 12	145	65	34	10	0			56
69	HSL 9	142	103	94	23	0	85.50	1	42
	HSL 12	223	69	23	5	2			74
	LSL 9	100	21	83	21	1			43
	LSL 12	56	25	160	12	1	90.52		63
73	HSL 9	65	40	218	38	1		3	62
	HSL 12	59	22	225	15	1			60
	LSL 9	115	59	11	43	0			29
	LSL 12	60	77	16	99	2	150.56		27
83	HSL 9	44	166	45	106	1		2	45
	HSL 12	80	83	19	139	1			27
	LSL 9	137	8	74	8	0			57
	LSL 12	100	16	124	13	1	162.31		36
94	HSL 9	51	47	240	23	1		1	12
	HSL 12	161	19	136	6	0			52

TABLE 5.

KILANDER HEALTH KNOWLEDGE TEST ITEMS WITH COMMON HIGH AND LOW DISCRIMINATION INDEX ACROSS FOUR SUBJECT GROUPS

High Discriminant Index	Groups				Low Discriminant Index	Groups			
Test Item	LSL 9	LSL 12	HSL 9	HSL 12	Test Item	LSL 9	LSL 12	HSL 9	HSL 12
73	59	65	62	56	45	-7	2	3	3
72	52	66	55	60	11	8	14	5	6
76	53	51	55	59	91	9	4	-1	13
78	66	66	50	64	17	5	7	13	-11

TABLE 6.
CORRELATION COEFFICIENTS AND REGRESSION COEFFICIENTS
OF TOTAL SCORE TO NINE CATEGORY SCORES
BY FOUR CRITERION GROUPS

Category of Test		LSL-9 (N=228) Category to total	LSL-12 (N=254) Category to total	HSL-9 (N=362) Category to total	HSL-12 (N=322) Category to total
Personal	r	.90	.91	.90	.94
Health	Slope	.33	.34	.34	.36
Nutrition	r	.75	.76	.67	.75
	Slope	.12	.12	.11	.11
Community	r	.80	.77	.73	.78
Health	Slope	.15	.13	.14	.12
Consumer	r	.78	.80	.80	.83
Health	Slope	.12	.13	.13	.12
Mental	r	.35	.43	.39	.55
Health	Slope	.03	.03	.03	.04
Family	r	.68	.74	.65	.70
Living	Slope	.09	.10	.09	.09
Stimulants &	r	.57	.66	.60	.77
Depressants	Slope	.07	.08	.08	.08
Safety	r	.65	.43	.44	.57
Education	Slope	.05	.03	.03	.03
First	r	.40	.54	.45	.64
Aid	Slope	.03	.04	.04	.04

TABLE 7.

BISERIAL CORRELATION COEFFICIENTS BY FOUR CRITERION GROUPS FOR SELECTED CATEGORY ITEMS OF KILANDER TEST

Category of item	Test Item	LSL-9 (N=228)		LSL-12 (N=254)		HSL-9 (N=362)		HSL-12 (N=322)	
		r Item to total	r Item to category	r Item to total	r Item to category	r Item to total	r Item to category	r Item to total	r Item to category
Family Living	4	-.06	-.09	-.06	-.24	-.11	-.26	-.18	-.29
	13	.12	.23	.15	.08	.12	.12	.19	.12
	14	.05	.16	.11	-.12	.07	.08	.06	.14
	19	.07	.33	.16	.29	.15	.36	.29	.47
	32	.29	.44	.34	.39	.28	.41	.19	.35
	37	.13	.14	.13	.07	.16	.10	.03	.06
	61	.55	.55	.48	.48	.43	.47	.44	.46
	64	.50	.45	.37	.42	.18	.33	.48	.42
	74	.51	.40	.29	.12	-.12	-.04	.42	.25
Mental Health	7	-.08	.14	-.05	.11	.01	.12	.05	.25
	43	-.01	-.26	-.13	-.30	-.09	-.30	.01	-.36
	50	.13	.14	-.08	-.05	-.02	-.11	-.11	-.22
	56	.33	.30	.16	-.27	-.01	.22	.25	.28
	91	.37	.20	.12	.23	-.04	.30	.21	.27
Safety Ed	27	.27	.57	-.22	.57	.20	.54	.29	.52
	66	.51	.43	.15	-.35	.01	.24	.27	.40
	99	.42	.37	-.34	.46	.26	.31	.52	.56

REFERENCES

1. Berg, I.A., Response Bias and Personality: The Deviation Hypothesis,
   Journal of Psychology, 40, 60-71, July 1955.

2. Berg, I.A., Deviant Responses and Deviant People: The Formulation
   of the Deviation Hypothesis, Journal of Counseling Psychology, 4,
   154-161, Summer 1957.

3. Bridges, A.F., A Valid Health Knowledge Test for College Freshmen,
   Doctoral dissertation, Indiana University, Bloomington, 1952.

4. Cambron, E.F., An Analytical Study of the Health Knowledge of Students
   in North Texas State College, doctoral dissertation, University of
   Texas, Austin, 1951.

5. Campbell, D.E. and Early, R.G., Comparison of Health Knowledge of
   Young Adults and Their Parents, Research Quarterly, 40:676-81,
   December 1969.

6. Campbell, D.E. and Foster, R.A., Health Knowledge of Young Adults
   from Two Socioeconomic Levels, Research Quarterly, In Press.

7. Coleman, A.E., Burkhardt, J.E. and Highfell, M.S., Comparison of
   Health Knowledge of Young Adult Under Achievers and Their Parents,
   Personal Communication, 1971.

8. Dearborn, T.H., Personal Health Knowledge of College Students Before
   Instruction, Research Quarterly, 29:154-59, May 1958.

9. Dixon, N.R., Social Class and Education: An Annotated Bibliography,
   The Harvard Educational Review, 23:330-38, Fall 1953.

10. Dees, M.H., Status of Health Education in Selected Texas Secondary
    Schools in 1966, Doctoral dissertation, University of Texas, Austin,
    1968.

11. Flanagan, J.C., General Considerations in the Selection of Test
    Items and a Short Method of Estimating the Product-moment Coefficient
    from the Tails of the Distribution, Journal of Educational Psychology,
    30:674-680, 1939.

12. Kilander, H.F., Health Knowledge of High School and College Students,
    Research Quarterly, 8:3-32, October 1937.

13. Mosley, J.W., An Analyses of Certain Factors Related to Health Know-
    ledge and Health Habits of College Freshman Women, Doctoral dissertation,
    Pennsylvania State College, 1950.

14. Murphy, M.A., Gain in Health Knowledge of Two Groups of Women Students
    Classified in Physical Education, Research Quarterly, 8:78-88, December
    1937.

15. Trice, E.X., A Study of the Health Knowledge of College Students in Louisiana, Master's thesis, Louisiana State University, Baton Rouge, 1946.

16. Warner, W.L., Meek, M., and Eells, K., Social Class in America, Chicago: Science Research Associates, 1949.