PREDICTION OF THE PERMANENCE OF INTERESTS

by

JAMES LUKE SMITH JR.

B. S., Kansas State College
of Agriculture and Applied Science, 1955

A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1956

TABLE OF CONTENTS

INTRODUCTION

The Strong Vocational Interest Test is one of the most widely used psychometric devices in vocational counseling today. It is frequently administered as an aid in the resolution of educational and vocational problems of college-age youth.

Since its introduction, it has become increasingly popular, due in large part to the intensive research that Strong and other psychologists have conducted on and with this instrument. This research has not only highlighted the central role of interests in effective vocational planning, but has defined in detail the validity and usefulness of this truly remarkable test.

The test itself is easy to administer, requiring practically no supervision or equipment other than a test form, answer sheet, pencil and a desk. The results are presented in graphic form enabling both the counselor and the counselee to form an accurate impression of the present vocational interests of the latter at a glance.

The degree to which the use of the Strong is a valid procedure in resolving the educational and vocational problems of a given student is a direct function of the degree to which the Strong is valid for that student. Since the persistent efforts of Strong and others have demonstrated the general validity of this instrument, it is necessary to conclude that for most students the test results can be used with validity in vocational counseling.

On the other hand, Strong (37) pp. 257-258, has shown that the major change in interests between ages 15 and 55 occurs in the first 10 years of this span. Inasmuch as this instability of interests occurs during the age range most frequently encountered by counselors of college students, it

would be desirable for the counselor to have some method of determining whether or not the picture of vocational interests obtained for a particular student could be taken to be stable. It might well be that the efforts of the counselor should be directed in different directions depending upon whether or not he was dealing with an individual whose vocational interests were stable or unstable. The identification of the probable stability of the profile is thus crucial.

The review of literature which follows is designed to give a research perspective to this problem, and to provide an appropriate setting for the present study.

REVIEW OF THE LITERATURE

The question of the permanence of interests has attracted a good deal of attention. Two major facets of this research area are pertinent to this study: (a) To what extent are interests permanent?, and (b) Can stability or instability of interest patterns be predicted? Investigations aimed at these two questions are reviewed below. For the most part, this review is restricted to research on the Strong Vocational Interest Test since the present study is on that instrument and since it holds an unchallenged position of leadership among available interest measures. Particularly relevant research on other interest tests will be reviewed when appropriate, however.

The Extent of Permanence of Interests

In answering the question of the extent of the permanence of interests, several methods of assessment have been utilized. These can be grouped as (a) test-retest correlations, (b) letter-grade changes, (c) profile correla-

tions, and (d) clinically judged stability. The review which follows will consider each of these methods.

Test-Retest Correlations. One of the most common ways of testing stability has been by the correlation of test-retest scores. The procedure has been to test a group of subjects on the Strong, wait a period of time, and then re-test them. Product-moment correlations are then computed for each scale. Essentially the method provides information on how consistently individuals have maintained their relative position in the group on the various keys.

Among the first, and one of the most detailed, studies of this type, is that by Strong (35, 39). He administered his test to 285 seniors at Stanford University in 1927. Five years later he retested 223 of these individuals and obtained retest correlations for this group. He again tested a large number of these individuals in 1937 and in 1949. This enabled him to obtain test-retest correlations for 5, 10, and 22 year periods. The mean correlations were surprizingly high--.78 for 5 years, .77 for 10 years, and .64 for 22 years. This finding led Strong to the generalization that approximately the same rank order is maintained in occupational interest scores for intervals ranging between 1 and 22 years.

While this study presents a fairly complete follow-up on college seniors, it does not cover the groups which are of primary interest to the school counselor, namely high school students and college freshmen. Fortunately, however, the literature does contain experiments which do give an indication of the permanence of interests of these two groups as measured by test-retest correlations.

Two such studies done on a limited set of scales are those of Canning,

et al. (3) and Van Dusen (53). In the former, 64 high school boys were tested as sophomores and retested as seniors. Over this two year period, an average correlation of .57 was obtained for the seven occupations employed in the study. Van Dusen (53) obtained similar results for 76 University of Florida men who were first tested as freshmen and retested 35 months later. His median test-retest correlation for 5 scales was .62.

While these two experiments are based on a limited number of scales, they do give a rough indication of the test-retest stability of the two groups tested. A logical next question is "What is the test-retest stability for students who originally took the test as high school students, and were retested at some time during their first two years in college?"

A recent study by Stordahl (33) provides information on this question. In 1951 he retested 181 male sophomores at the University of Minnesota who had been tested originally as high school seniors in 1949. He divided his population into two groups--111 metropolitan and 70 non-metropolitan students. For these groups he obtained median test-retest correlations of .72 for the metropolitan sample and .67 for the non-metropolitan sample.

The studies cited above provide a fairly consistent picture of test-retest stability correlations on the Strong. The amount of generalization possible is limited by the fact that at one or both testing periods the subjects were students. Fortunately a study by Powers (25) sheds some light on the question of test-retest correlations for a group of gainfully employed individuals. She located a sample of 109 cases who were originally tested in 1931 and who were retested in 1941. At the first testing, the mean age of her subjects was 34. The blanks were scored on all 44 keys of the 1938 revision. Mean test-retest correlation for this sample was .69. While this is not as high as that obtained by Strong (35, 39) for his students 10 years

after testing (.77), it is sufficiently high to indicate that individuals in this group also tend to maintain their position within the group on the various interest keys.

These studies consistently suggest a high degree of stability in the rankings of individuals on the scales of the Strong test. Similar studies have been done with the Kuder Preference Record, although the bulk of this research has been with high school students. Two studies will serve to point out that, in general, the results are parallel to those cited above for the Strong.

In 1948, Reid (27) administered the Kuder Preference Record Form BB together with other tests to the entering freshman class at Westminister College, Pennsylvania. Of this class, 145 members volunteered to retake the test 15 months after the initial administration. There were 63 males and 82 females in the retest group and the mean age at the time of original testing was 18.4 years. He obtained test-retest correlations which ranged from .72 on computational interest to .84 for persuasive, with a median correlation of .77. This is somewhat higher than the results obtained by Van Dusen (53) on the Strong, but closely approximates the results obtained by Stordahl (33).

In an attempt to determine the test-retest correlations on the Kuder for high school students, Rosenberg (28), selected as his sample 91 boys and 86 girls who had graduated from East Syracuse High School during the years 1949, 1950, and 1951. He compared ninth grade and twelfth grade scores of these students and found a median correlation of .61 for both the boys and the girls. This correlation closely approximates that obtained by Canning, et al. (3) in their study of the Strong.

This section of the review of the literature has presented an illustrative group of studies aimed at the question of the degree to which interests

are permanent. They have in common the use of the test-retest correlation coefficient as the measure of permanence. Other studies could have been cited, but the results are so consistent as to be repetitious. In general, the weight of the data suggests that the interests of college and post-college groups are somewhat more permanent than those for high school students. However, interest stability (as measured by test-retest correlations) of all groups is high even over lengthy periods of time.

Correlations of the magnitude reported above are both impressive and misleading. In psychology, correlations above .60 are rare, regardless of what type of behavior is being predicted. On the other hand, a correlation of .71 between two variables indicates that only 50 per cent of the variance of one is accounted for by the other. Thus, there would appear to be a good amount of shifting in position within a group on interest scales. While the data require the generalization that interests remain surprizingly stable (according to this measurement method), they also require the generalization that there is a significant amount of instability in interests.

One final comment is in order before leaving this topic. It should be borne in mind in summarizing these studies, that this method of measuring stability tells only how well an individual maintained his relative position within the group. Theoretically it is possible to have a perfect correlation and still marked changes in the level of interests could occur. For example, all subjects might score "C" on a given scale during the first administration of the Strong, and "A" on the retest. If they maintained their relative standing on that scale, the correlation would be 1.00; but one would have to conclude that interests had shifted radically. For this reason it is important to look at the second method of measuring the stability of interests

which is letter grade changes.

Letter Grade Changes. Results of the Strong are given in terms of letter grades. Regarding the purpose of this system of reporting scores Strong states:

> The objective in using letter ratings of A, B, and C has been to distinguish between "Yes, you have the interests of the occupation" (A), and "No, you do not have those interests" (C), and to throw the doubtful ones into the B ratings (37, p. 370).

With such an unambiguous interpretation of letter grades available, studies of changes in letter grades from one test to the next are obviously relevant to the question of the permanence of interest.

Here again one finds that the first investigator to conduct such a study was Strong (37). In 1937 he retested 168 individuals who had originally been tested as college seniors in 1927. Scores were obtained on 18 keys. Regarding changes in letter grades, Strong concluded, in part:

> Of the 12 per cent of A ratings of seniors in 1927, ten years later, 7.4 per cent were A ratings, 2.0 per cent were B +, and 1.9 per cent were B ratings..... Evidently if a senior had an A rating there was very little chance of his receiving anything but an A, a B +, or a B rating a decade later. Furthermore, there were 34.0 per cent C ratings in 1927; in ten years these had changed so that there were 21.4 per cent C ratings, 5.5 per cent C + ratings, 4 per cent B- ratings, and 2.0 per cent B ratings. If a senior had a C rating in 1927, there are 2.3 chances in a hundred that it might be raised to a B + in 1937 and 0.9 chance in a hundred that it might be raised to an A rating (37, p. 366).

Having seen these results for college seniors we can now look at a similar study done with college freshmen. In 1950 Trinkhaus (51) retested 212 Yale Alumni who had originally been tested in 1935 and 1936 as college freshmen. His results, based on 39 occupational scores, led him to conclude that: "The data on changes in letter grades showed (a) the extreme scores to be most stable with the low scores being more stable than the high scores" (51, p. 646).

In like manner, Stordahl, with his 181 college sophomores who had been tested originally as high school seniors, was led to conclude:

> On the average C grades were the most stable, 68 per cent of the C grades on the first test being C grades on the second test. The second most stable letter grade was A with 60 per cent of the letter grades being identical on the test and retest. The intermediate letter grades were less stable. By combining the letter grades so that B included B +, B and B-, and C included C + and C, it was found that 73 per cent of the C grades on the test remained C grades on the retest. 59 per cent of the A and 59 per cent of the B grades remained constant (34, p. 425).

The major findings of these three studies are summarized in Table 1. In general, the results of Stordahl and Strong are in close agreement, while Trinkhaus' data are somewhat divergent. It would be interesting to speculate on factors related to this fact, but the number of uncontrolled variables (time interval, cultural differences, number of scales used) is so large as to make such an undertaking of dubious value.

Table 1. Test-retest stability of letter grades from three separate samples of students at Stanford, Yale, and the University of Minnesota.

Test Grade	Per cent Retest Grade		
	C	B	A
A			
Stanford	2.5	35.8	61.7
Yale	10.6	43.5	45.9
Minnesota	5.0	36.0	59.0
B			
Stanford	23.2	57.3	19.5
Yale	31.6	52.3	16.1
Minnesota	24.1	59.3	16.6
C			
Stanford	71.6	26.8	1.7
Yale	73.4	24.1	2.5
Minnesota	73.3	25.0	1.7

Stanford data are from Strong on 168 college seniors retested ten years later on 18 occupations.

Yale data are from Trinkhaus on 212 college freshmen retested 14-15 years later on 39 occupations.

Minnesota data are from Stordahl on 181 high school seniors retested 2-2.5 years later on 44 occupations.

In order to determine the consistency of letter grades for students who were tested and retested in high school, one may refer again to the study by Canning, et al. (3). They found that their percentage (64%) of identical letter grades received by high school boys retested two years later was almost identical to the percentage (63.3%) of identical ratings that were received by Strong's college seniors retested after a five-year period.

Again, in this section, an attempt has been made to show the stability of interests of the four groups with which the student counselor would be most concerned--college seniors, college freshmen, high school students retested sometime during their first two college years, and high school students. Once again the evidence points conclusively to the fact that stability of interests, as measured by the number of identical letter grades received on test and retest, is surprisingly high on the average. The evidence further points out that the C grades are the most stable, A grades next, and B grades the least stable. However, by lumping the grades together so that C includes C + and B includes B- and B +, the discrepancy in stability between the B and A grades tends to disappear.

The generalization regarding letter grades is somewhat misleading. "B" ratings include standard scores from 30 to 44 inclusive. "A" and "C" ratings on the other hand, include a much greater range of standard scores. It is possible to have a discrepancy of 30 standard score points on two testings and still obtain a pair of "C" or a pair of "A" ratings.

Despite the high percentages of identical letter grade ratings, it is again necessary to exercise caution in generalizing. The highest percentage of identical ratings obtained (73.3) clearly indicates that there is a sizeable group of individuals whose interests are as yet not stable according

to this method of assessing interest stability.

This section on letter grade changes and the preceding one on test-retest correlations have dealt with how closely the test and retest results on a single scale have agreed. The latter was concerned with the consistency with which individuals could be ranked on a given scale, while the former provided information on the consistency of absolute (letter grade) scores. If one turned his attention to the stability of test scores within a given individual he would be asking how closely the entire profile agreed on two or more testings. This question focuses attention on the third method of measuring the stability of vocational interests, namely profile correlations.

Profile Correlations. The procedure in this case has been to rank the sigma scores obtained from each individual's profile on both test and retest and then to correlate the ranked scores. This method appears to have more relevance than the previous two to the vocational counselor, since his concern with interest tests is with the entire pattern of scores obtained by the individual rather than with any single score.

As in the other two cases one again finds that the first to investigate this aspect of stability was Strong (43). For the seniors at Stanford who were first tested in 1927, he found a median rho on profiles of .86 after 5 years; .82 after 10 years; and .76 after 22 years.

If attention is turned to college freshmen, again it is found that Strong (37, 43) has pioneered the way. In 1930, 306 Stanford University freshmen filled out the Strong Vocational Interest Blank. Many of these students filled out the blank in 1931 and 1939 and 202 did so in 1949. From these blanks he has computed the profile stability for college freshmen for periods of 1, 9, and 19 years. His median rhos in this experiment were:

.88 for 1 year; .67 for 9 years; and .72 for 19 years. These relatively high mean rhos are again interpreted to mean that the vocational interests of college freshmen are surprizingly stable.

Relevant data on high school students has been published by Taylor and Carter (46). They administered the Strong to 58 high school juniors, and retested them a year later. Their rho coefficients ranged from .65 to .99 with a median of .74. This is not as high as that obtained by Strong for his college freshmen after one year (.88). However, the fact that they scored only 12 scales whereas Strong scored 20 to 34 scales in his study may account for this difference.

For similar data on high school students who were retested during their first two years in college, reference can again be made to Stordahl (33). He found that the median rho was .74 for the metropolitan group and .72 for the non-metropolitan group.

These studies give a fairly clear picture of the "within individual" stability of interests for various student groups. Similar information for a more heterogeneous group of individuals is provided by Powers (25). It will be recalled that she had Strong tests taken in 1931 and 1941 by a group of 109 gainfully employed men. For this group the median rho for 44 scales was .80 which is almost identical with that obtained by Strong (.82) for his seniors who were retested 10 years later. Table 2 has been prepared as a summary of the profile correlation data cited above.

These data point out fairly conclusively that for most of the subjects tested, the <u>within individual</u> stability of vocational interests is quite high. Once again, however, caution should be taken in generalizing. While median rhos in the neighborhood of .70 to .80 are impressive, the median may conceal as much as it reveals. Half of the cases fell below this summary

Table 2. Permanence of interest profiles as obtained by four investigators.

Investigator	Level	Time Interval (years)	Median Rho
Strong	College Seniors	5	.86
	College Seniors	10	.82
	College Seniors	22	.76
	College Freshmen	1	.88
	College Freshmen	9	.67
	College Freshmen	19	.72
Taylor and Carter	High School Juniors	1	.74
Stordahl	High School Seniors	2	.74 & .72*
Powers	Gainfully Employed	10	.80

*.74 for metropolitan, .72 for non-metropolitan.

figure. It seems likely that for a substantial number of these cases, the profile correlations are so low as to indicate instability in interest pattern.

This raises the question of how high a rho must be before it is indicative of stability in interests. A related question may be derived by the following reasoning. It is possible that two profiles would correlate 1.00 and still a counselor, in interpreting the profile would find they had different meanings. For example, profile 1 may contain nothing but B and B- scores, while profile 2 may contain half A's and half C's. If the rank order of the scores was identical, the profiles would correlate 1.00, but a counselor would surely say different things about the two. This fact, plus the point made in the preceding paragraph, raises the question of whether or not rank-order correlations between profiles is a meaningful way of expressing stability. Data

from our fourth method of measuring interest stability, namely clinically judged stability, is relevant here.

Clinically Judged Stability. Despite the fact that counselors tend to pride themselves on their ability to interpret "dynamic patterns" rather than "static scores" on interest tests, there has been very little research effort devoted to the question of the stability of these clinically recognizeable patterns. Only three studies were found which gave attention to this problem.

Hoyt (17) selected 72 pairs of profiles from Stordahl's original data. Eight of these pairs correlated .50 to .54; 8 other pairs correlated .55 to .59; and so on. The final 8 pairs correlated .90 to .94. Three experienced counselors then rated each pair of profiles as to how much different the interpretation of the first profile would be from the second profile. A five point scale which ranged from "1 - no difference" to "3 - many features similar, but at least one important difference" to "5 - definitely different interpretations" was used. Uniformly high judge agreement was obtained, with at least 2 out of the 3 counselors agreeing on 97 per cent of the ratings. No discrepancy larger than 2 scale points was found. The average counselor rating obtained for each set of profiles together with the corresponding rank-order rhos is shown in Table 3.

Perusal of this table provides an answer to the question as to whether or not rank-order correlations are a meaningful way of expressing stability. It can be seen that there is an obvious relationship between "statistical stability" and "clinical stability". In fact the rank order correlation between the two sets of data was .95. Darley and Hagenah summarize these results aptly:

> When the rank order between the individual's first and second testing reaches or exceeds .75, experienced counselors make essentially

Table 3. Relationship of counselor judgment to the statistical stability of interest profiles.*

Rho Between Profiles	Average Counselor Rating
.90 - .94	1.63
.85 - .89	1.88
.80 - .84	2.38
.75 - .79	2.21
.70 - .74	2.54
.65 - .69	2.63
.60 - .64	3.04
.55 - .59	3.17
.50 - .54	3.04

*From Hoyt (18, p. 10).

> the same interpretation of the two profiles. If a test-retest correlation ranges from .65 to .75, the two testings show at least one important difference in interpretation. Test-retest rank order values below .65 may result in considerable differences in interpretation and represent major changes in interest patterns (6, p. 45).

This then provides a comparison between the rank-order rhos and clinically judged stability. Our primary concern, however, is the proportion of individuals whose profiles would represent unstable interests. In this study Hoyt found that the following percentages of profiles had rank order rhos of .65 or less, thereby indicating some major change in the clinical interpretation of them:

1. 17 per cent of the college sophomore - college senior profiles.
2. 40 per cent of the high school senior - college sophomore profiles.
3. 47 per cent of the high school senior - college senior profiles.

Strong (43) p. 77, has made two rough estimates pertaining to this same question. In one case he studied the changes in letter grades of 9510 scores on 15 scales and found that in roughly 25 per cent of instances the shift in letter grade was radical enough to induce a counselor to interpret the scores differently. In his second rough estimate, he clinically inspected each pair of test-retest scores on 34 scales for the first 10 freshmen in his files and found that 11 per cent of the changes in scores did affect the interpretation of the profile.

While these two estimates do give an indication of instability of interests when the profiles are clinically judged, they indicate only the per cent of specific scores which were unstable, not the number of profiles. In addition, in the second estimate the sample was too small to provide any conclusive evidence.

The three studies cited above dealt with the Strong Vocational Interest Test. A recent study by McCoy (24) has been done on the same question using the Kuder Preference Record.

In 1954, he had 3 experienced judges clinically classify 184 pairs (101 females and 83 males) of test-retest Kuder profiles into three groups--"stable", "change" and "doubtful". The two sexes were treated separately. The educational status of the students ranged from grade 8 to grade 10 at the time of the initial test and grade 11 to grade 12 at the time of retest. The time between test and retest was 24 to 42 months with a mean of 33 months. After three clinical inspections with a composite agreement among the three judges, he found that 66.3 per cent of the males and 60.1 per cent of the females fell into the "change" category. These percentages were somewhat higher than those obtained by Hoyt (17) in his experiment. Whether this

discrepancy is due to the fact that McCoy used a younger sample, or to the difference between the Kuder and the Strong, or to the method of determining stability cannot be ascertained without further experimentation.

Although these three studies show that interests are judged to be stable for the majority of college age individuals (and slightly less than the majority of high school students), a substantial proportion seem to have unstable interests. Just how big that proportion is undoubtedly is a function of (a) when the test was first administered and (b) the time interval before the retest. Thus estimates vary from as low as 17 per cent to as high as 66 per cent with unstable patterns. Assuming the true figure lies somewhere in that range, it seems safe to conclude that interest stability represents an important individual difference. The ability to predict this characteristic would appear to be highly desirable.

Thus far, the effort in this review has been devoted to answering the first question posed, namely the extent to which interests are stable. The fact that interests are reasonably stable for the majority of individuals has been stressed. At the same time, it has been pointed out consistently that for a substantial number of individuals, this generalization does not hold.

The desirability of identifying these individuals at the time they are tested introduces the second question with which the literature review is concerned: Can stability or instability of interest patterns be predicted?

Attempts to Devise Stability Scales

Although a considerable amount of work has been done on determining the extent of interest stability, very few investigators have attempted to devise

methods whereby stability could be predicted.

Again the pioneering work of E. K. Strong (37) must be recognized. In about 1933 he published his original "Interest Maturity Scale" which he defined as the "quantitative measurement of change of interests with age.... the degree to which one has interests of 55-year old men as compared with those of 15-year-old boys" (37, p. 247). One criterion group used in devising this scale was composed of a representative group of 472 15-year old boys. These 472 boys came from 18 school systems in California, 3 in Oregon and 1 in Washington. The grade composition of this group is shown in Table 4.

Table 4. Composition of 15-year-old criterion group, according to school grades, in comparison with distribution of 15-year-olds in schools of California.*

School Grade	15-Year-Old School Children In California		15-Year-Old Criterion Group	
	Total	Percentage	Total	Percentage
Grade 7	4,302	6.83	31	6.5
Grade 8	9,703	15.40	73	15.5
Grade 9	18,674	29.64	140	29.7
Grade 10	19,944	31.65	150	31.8
Grade 11	8,844	14.04	66	14.0
Grade 12	1,539	2.44	12	2.5
Total	63,006	100.00	472	100.0
Average School Grade	9.38		9.39	

*Adapted from Strong (37, p. 715).

A second criterion group was that of 55-year old men employed in a variety of occupations. Table 5 gives the occupational classification of these men.

Table 5. Occupational composition of 55-year-old group for original IM scale.*

	N		N
Group I:[a]		Group III a:	
Architect	17	Minister	50
Artist	26	Teacher	33
Engineer[b]	30		
Mathematician	30	Total	83
Physician[c]	38		
Physicist	30	Group III b:	
Psychologist	30	City School Superintendent	30
Miscellaneous	12	Personnel Manager	21
		Y.M.C.A.	17
Total	213		
		Total	68
Group II a:			
Advertising man	3	Group IV:	
Author	20	Purchasing Agent	11
Lawyer	34	Miscellaneous	24
Newspaper editor	36		
		Total	35
Total	93		
		Group V:	
Group II b:		Certified Public Accountant	13
Life Insurance Salesman	24	Miscellaneous:	
Real Estate Salesman	9	Hotel Manager	17
Miscellaneous	14	Business Manager	15
		Miscellaneous	48
Total	47		
		Total	80
		Grand Total	632

*From Strong (37, Table 192, p. 716).
[a]Original grouping of occupations used at that time.
[b]Actually based on 27 mining, 25 civil, 23 electrical, and 25 mechanical engineers, but reduced proportionately to equal 30.
[c]Actually based on 76 and reduced to equal 38.

Strong's test was administered to both criterion groups. The responses of each group were compared on each item of the test, and differentiating items were retained to form the original "Interest Maturity" scale.

Strong was discontented with the composition of the 55-year-old group. It represented primarily successful men in professional fields of work. Little

or no representation was allotted to executives and skilled or unskilled laborers. The revised Interest Maturity Scale was prepared in an attempt to free such a scale from undue influence of educational and occupational attainment. At the time of this revision, a second refinement was introduced. Strong had assumed in his original work that liking for an activity increased or decreased progressively from 15 to 55 years. Subsequent investigation had failed to substantiate this assumption for many items. He therefore included samples of men at ages 25, 35, and 45 in addition to the 15 and 55 year old groups. In this way, he sought to determine whether changes in interest for each item increased or decreased progressively from 15 to 55 years or whether these changes progressed to some intermediate age and remained constant or reversed themselves thereafter.

He used as his sample for the revised scale (a) the original criterion groups at ages 15 and 55; (b) 80 pairs of fathers and sons who averaged 58 and 22 years of age; (c) 4 groups of different ages (25, 35, 45, and 55) who were representative of occupations at those ages as listed by the United States Census; (d) Stanford Seniors tested in 1927, and retested in 1932 and 1937; (e) High school seniors tested in 1927 and retested 6 years later; (f) Stanford graduate students who were tested in 1927 and 1932 (average ages 25 and 30); and (g) Stanford graduate students tested in 1927 and retested in 1932 (average ages 30 and 35). A total of 5,000 males were used, 894 representative of the population at large and 4,106 students. Responses to each Strong item by members of each of the age groups--15, 25, 35, 45, and 55 were tabulated. This tabulation revealed that there was a characteristic change in interests from age 15 to age 25. After age 25, the per cent of each age group liking or disliking various activities remained fairly constant, al-

though there were some reversals from the 15-25 year old trends.

On the basis of this information, Strong constructed a revised Interest Maturity Scale using 15-year-olds and 25-year-olds as comparison groups. This scale was meant to apply only to men in the age range of 15 to 25, and primarily to men under 21, since the greatest shift in interests comes in the first 5 years of that span.

After constructing this scale, Strong (37) then attempted to ascertain whether or not a high Interest Maturity score would be an indication of interest stability. In this investigation he used three groups of high school juniors. One group contained 10 individuals with the lowest Interest Maturity scores; another group contained 10 individuals with an Interest Maturity score equal to the median of the entire group; and the third group contained 10 individuals who had the 10 highest Interest Maturity scores. Tests were administered at ages 16.5 and at 22.5 years and the mean Interest Maturity scores obtained for each group was respectively: 28 and 45, 45 and 51, and 60 and 51. The rank order correlations between profiles for each was .68, .79, and .91. These results led Strong to conclude: "Evidently the higher the IM score at 16.5 years the less chance there is for occupational scores to change in the next few years and the more likely it is that such scores will agree on both occasions" (37, p. 281).

Recently, however, Stordahl (33) in his more extensive study cited earlier has cast considerable doubt on the validity of the Interest Maturity Scale as a predictor of stability of interests. It will be recalled that he obtained rank-order correlations on all 44 occupations for each pair of profiles, using a two year test-retest period. On the basis of these correlations, he divided the sample into "High", "Average", and "Low" stability groups. Table 6 shows the mean Interest Maturity Score and the standard

deviation for each group as obtained by Stordahl:

Table 6.* Mean interest maturity scores and standard deviations for students with high, average, and low interest stability.

Group	N	Mean	S.D.
High	60	46.7	8.8
Average	61	48.6	7.0
Low	60	46.5	8.9

*Adapted from Stordahl.

The small differences in mean score on the Interest Maturity Scale were not statistically significant, leading Stordahl to conclude that "these results fail to substantiate the assumption of a positive relationship between interest stability and interest maturity" (33, p. 341).

As a result of these conflicting findings between Strong and Stordahl, Hoyt (17) added to Stordahl's data and analyzed it in more detail. In his study he asked four basic questions. The first one was whether or not Stordahl's findings were a function of the fact that his subjects were quite young at the time of retest. Strong's study of high school juniors had utilized a six-year interval between tests, whereas Stordahl had only a two-year interval. In order to help answer this question and succeeding ones, Hoyt retested as college seniors 121 of Stordahl's original 181 students. He then computed rank order correlations between the 1949 and 1953 profiles of these 121 students. On the basis of these correlations, the sample was divided into high, average and low stability groups, and the mean Interest Maturity Score for each group were computed. The results of this analysis are shown in Table 7.

Table 7.* Relationship between interest maturity scores and interest stability over a four-year period.

Stability Group	Rho between 1949 and 1953 profiles: N	Range	Mean**	IM Score: Mean	S.D.
High	38	.77 - .99	.86	46.1	9.3
Average	45	.55 - .76	.67	48.6	6.7
Low	38	-.48 - .54	.28	45.3	9.4

*From Hoyt (18, p. 4).
**After transforming ρ's to z's.

An analysis of variance resulted in acceptance of the null hypothesis. Increasing the time interval between tests, and thus increasing the probability that the final test revealed stable interests, did not resolve the Strong-Stordahl differences.

A second question posed in Hoyt's study was, "Are the differences in the Stordahl and Strong findings due to different methods of analysis?" It will be recalled that Strong (37) used groups with different Interest Maturity Scores and studied the differences in profile stability of these groups, whereas Stordahl used groups with different profile stability and studied the differences in Interest Maturity Scores. In order to answer this question, Stordahl's data were re-analyzed using groups with the highest and the lowest Interest Maturity Scores as a starting point. Profiles from the 1949-1953 interval were treated in like manner. The differences in profile rhos were so slight that they were of neither statistical nor practical significance. Thus the method of analysis did not appear to be a determinant of the Strong-Stordahl discrepancy.

A third question investigated by Hoyt was, "Can the sensitivity of the

Interest Maturity scale as a predictor of interest stability be increased by statistically controlling its occupational saturation?"

It might be well to note at this point that Strong's criterion groups of Y.M.C.A. secretaries, personnel managers, social science teachers, ministers, and accountants averaged 55 or higher on Interest Maturity, whereas dentists, physicians, mathematicians, and journalists all averaged less than 50.

These data suggest that an Interest Maturity score of 45 may have different meanings for individuals who made A scores on mathematician and those who made A scores on minister. It is this correlation between occupational scores and the Interest Maturity score which is referred to as an "occupational saturation." Hoyt attempted to control the occupational saturation by sorting the profiles according to primary patterns (Darley and Hagenah, 6). All students with a primary pattern in a single group were combined, and the correlations between Interest Maturity and Profile Stability (1949 and 1951) for 6 interest groups were computed. The highest correlation obtained (.27) was not significantly different from zero, leading to the conclusion that controlling for occupational saturation did not increase the sensitivity of the IM scale as a predictor of interest stability.

A final question posed by Hoyt (17), and one particularly relevant to the present investigation, was: "If the Interest Maturity scale does not predict the stability of interest patterns, can an empirical scale be devised which will?"

In this part of the study the 1949 responses of each of the three stability groups (high, average, and low) defined by Stordahl were compared. In an effort to purify the criterion, the 10 students with the most marginal stability scores from both the high and the low groups were eliminated. This

then left a total of 161 students, 50 each in the high and the low groups and 61 in the average group. Tabulations were made of the number and proportion of students in each group who answered "L", "I", and "D" to each of the 400 Strong items. For those items in which the proportion of the average group fell in between the proportions for the high and low group, tests of significance were run. Using the 5 per cent probability level, 37 items were obtained which met the criteria of "(a) statistically differentiating the high and low groups, and, (b) having the proportion of the average group fall in between the proportions for the two extreme groups."

Cross-validation of this scale was attempted on three different samples. The first were the sophomores of Stordahl's group who were retested as seniors. This key correlated only .15 with the measure of stability used (z'). A second sample was provided by Powers (25) and described above. Here a correlation of .27 was obtained with z' which was statistically, but not practically, significant. The final attempt at cross-validation was with a group of 21 University high school boys tested as sophomores and again as seniors in high school. Here the correlation between z' scores and the new key was only .04. These results led Hoyt to conclude that the new key lacked sufficient validity to warrant further investigation.

We have seen from this section that the two major efforts to devise a scale which would predict stability of interests using the Strong Vocational Interest Blank have met with failure. Have the results been more gratifying with other interest measures?

Actually only one major study concerning the construction of a stability scale for other interest tests was found. This was a recent investigation of the Kuder Preference Record by McCoy (24) and referred to earlier.

In 1954, using the Kuder Preference Record (Vocational) Form C, and

treating the sexes as separated groups, he attempted to derive a stability scale. It will be recalled that he used 184 pairs of tests (101 females, and 83 males) from communities and states surrounding the University of Missouri. Each profile was coded, and experienced judges divided them into three groups--"change", "stable", or "doubtful". He then examined the responses of the "change" and "stable" groups to each of the Kuder items. Items which differentiated the "change" group from the "stable" group at the 10 per cent level of confidence were retained to form a new scale for predicting stability on the Kuder. For the boys 84 items met this criterion, while 89 items were similarly chosen for the girls.

In order to test the validity of these scales, they were applied first to the tests from which they had been derived in a preliminary "circular type of validation". Naturally, in this case the scales did differentiate between the "change" and "stable" groups for both boys and girls.

He then obtained a new sample of high school students who had taken the test twice with two to three years intervening between test and retest. Again, "change", "stable", and "doubtful" groups were formed. When applied to these new samples, McCoy's stability scale failed to satisfactorily discriminate between the "change" and "stable" groups for either males or females.

In this section, an attempt has been made to answer the second question which was proposed, namely, "Can interest stability or instability be predicted?" Only a few attempts to devise a method of predicting stability have been made. Of these, the record is consistent. No adequate indicator of interest stability has been discovered.

The argument that such a predictor is not needed because interests are,

for the most part, stable characteristics of individuals does not seem tenable. All studies of the prevalence of stable interests indicate that for a sizeable group of adolescents, interest measures are not sufficiently stable to warrant their serious use in vocational counseling.

Just how big this group is must remain problematical for the moment, but the studies of Strong, Hoyt and McCoy suggest a minimum figure of 10 per cent and a maximum of near 60 per cent. Even if these students represent a minority of the population, they are of particular interest to the counselor because, for them, the procedure of using the Strong Vocational Interest Test as an aid in the resolution of their vocational and educational problems may not be a valid one.

It is obvious that if some method were devised whereby these individuals could be identified at the time the test was given, the advantages to both the counselor and the counselee would be immeasurable. A successful scale for determining this has not yet been devised.

This study was undertaken in an attempt to shed some light on this important problem.

EXPERIMENTAL DESIGN

Hypotheses

Three hypotheses were proposed for this study. They were:

1. A stable Strong Vocational Interest Profile (one that is reliable over time) will have a higher order of consistency than an unstable one. This consistency can be defined as:

   a. The relationship between the subscales that the individual has high scores on.

b. The relationship between the subscales that the individual has low scores on.

c. The relationship between the subscales that the individual has high scores on and those subscales that the individual has low scores on.

2. Consistency of response to a core of items will be related to profile stability.

3. A count of the "likes" on the items that Strong has shown a characteristic increase in liking and decrease in disliking from age 15 to 25 will produce a valid index of stability.

The first hypothesis was suggested in personal correspondence by Dr. Seymour Levy. In more detailed form the reasoning behind this hypothesis is as follows. If an individual scores high on a given occupation, say Dentist, we know from Strong's table of intercorrelations that he could be expected to score high on certain other occupations (Physician, Artist, Architect) and low on others (Sales Manager, Real Estate Salesman). Now if an unexpected pattern occurs (high Dentist, low Physician, high Sales Manager) it seems reasonable to suppose that one explanation for this would be that the individual's interests are not really stable as yet. In a way, the hypothesis is that stability is related to consistency, and that over a period of time the individual simultaneously becomes more stable and more consistent in his interests. Although the point is not germane to the test of the hypothesis, it is assumed, theoretically, that a necessary condition for stability is consistency. Further, it is assumed that a meaningful measure of consistency can be obtained by determining whether an "expected" or an "unexpected" pattern of interests appears.

In the case of the second hypothesis, it was felt that if a core interest

concept could be identified and if items pertaining to that particular concept were analyzed, it would be found that a stable profile would be characterized by consistent answers to these questions, whereas an unstable profile would not. In other words, for any particular core group of items, individuals with stable interests will respond consistently, whether that response is "Like", "Indifferent", or "Dislike". The second hypothesis is thus much like the first, except it is applied to the relationships among items instead of that among scales.

The third hypothesis was suggested by Hoyt (17). Using Strong's finding that certain items show a characteristic increase in liking and a decrease in disliking from age 15 to 25, he proposed that a simple count of the number of likes or dislikes among these items might provide an index of profile stability. The reasoning behind this hypothesis, of course, is that individuals with stable profiles would resemble Strong's older group in specific preferences more closely than would individuals with unstable interests.

Sample

A total of 176 of the 181 students used by Stordahl (33) were chosen as the basic sample for this study. Five of the students were found to have z' scores of either a negative value or zero and they were eliminated in order to simplify computational problems.

As was pointed out in the literature review, Stordahl administered the Strong test in 1951 to 181 sophomores at the University of Minnesota who had originally taken it as high school seniors in 1949. He computed coefficients of concordance for each individual's test and retest scores and arbitrarily divided the profiles into three groups of approximately equal size on the basis of their W values.

Those with coefficients of concordance of .906 to .977 were designated as the "high" stability group (N-60), those with values of .820 to .905 were designated as the "average" stability group (N-61), and those with values of .419 to .818 were designated as the "low" stability group (N-60).

In 1953 Hoyt (17) administered the Strong for a third time to 121 of these students. For a part of the present study, 116 of these students constituted the sample. Again, five were eliminated in order to facilitate computations.

Predictive Indices

As noted earlier, the first hypothesis stated that profile stability was a function of:

1. The extent to which scales on which the individual scored high were intercorrelated.

2. The extent to which scales on which the individual scored low on were intercorrelated, and;

3. The extent to which scales on which the individual scored high were correlated with the scales on which he scored low.

A modification of Levy's proposed method of testing these sub-hypotheses was used. First a form was prepared in which weights were assigned to every pair of occupations in accordance with the degree of intercorrelation between the scales. Table 8 gives the schedule of weights which was used.

The form thus prepared is included in the Appendix. It will be noted that this form does not include the newer scales (Psychologist (Rev), Veterinarian, Industrial Arts Teacher, Vocational Agriculture Teacher, Public Administrator, Senior CPA, Mortician, and Pharmacist) as no correlations for them were listed in Strong's table (37, Table 193, p. 716).

Table 8. Schedule of weights assigned and corresponding correlations.

Weight	Correlation Between Occupational Scores
4	.80 and higher
3	.60 to .79
2	.40 to .59
1	.20 to .39
0	-.19 to -.19
-1	-.20 to -.39
-2	-.40 to -.59
-3	-.60 to -.79
-4	-.80 and higher

Using this matrix of weights, indices designed to test sub-hypotheses a, b, and c (and thus hypothesis 1) were computed in the following manner: First the five highest and the five lowest scales were identified. The "Highs" index was computed by considering only the five highest scores. A total of 10 values were obtained from the matrix described above, 1 for each of the 10 possible comparisons within the 5 highest occupations. These values were summed to yield the "Highs" index.

Similarly the "Lows" index was computed by considering only the five lowest scores. Again the 10 obtainable weights were summed to yield the "Lows" score.

For the "Highs versus Lows" score, each high scale was compared, individually, with all five low scales. Thus 25 scale values (1 for each of the 25 possible comparisons) were obtained. These values were then summed to yield the "Highs versus Lows" score. An example of the scoring method follows:

Scores obtained in 1949 by Subject #64

Highest Scores	Lowest Scores
1. Purchasing Agent	1. Minister
2. Office Man	2. City School Superintendent
3. Farmer	3. Personnel Director
4. Math, Phys. Sci. Teacher	4. Osteopath
5. Banker	5. Physician

Computation of "Highs" Index

Comparison	Weight
1 vs 2	3
1 vs 3	0
1 vs 4	0
1 vs 5	2
2 vs 3	0
2 vs 4	1
2 vs 5	3
3 vs 4	3
3 vs 5	0
4 vs 5	-1
Score: "Highs"	11

Computation of "Lows" Index

Comparison	Weight
1 vs 2	3
1 vs 3	2
1 vs 4	1
1 vs 5	0
2 vs 3	2
2 vs 4	0
2 vs 5	-1
3 vs 4	0
3 vs 5	-2
4 vs 5	3
"Lows"	8

Computation of "Highs versus Lows" Index

Comparison	Weight
High 1 vs Low 1	3
High 1 vs Low 2	-1
High 1 vs Low 3	0
High 1 vs Low 4	-2
High 1 vs Low 5	-2
High 2 vs Low 1	0
High 2 vs Low 2	0
High 2 vs Low 3	2
High 2 vs Low 4	-1
High 2 vs Low 5	-3
High 3 vs Low 1	-1
High 3 vs Low 2	-2
High 3 vs Low 3	-2
High 3 vs Low 4	1
High 3 vs Low 5	1
High 4 vs Low 1	1
High 4 vs Low 2	0
High 4 vs Low 3	0
High 4 vs Low 4	2
High 4 vs Low 5	0
High 5 vs Low 1	-1

High 5 vs Low 2	0
High 5 vs Low 3	0
High 5 vs Low 4	-2
High 5 vs Low 5	-3
Score: "Highs versus Lows"	-10

In testing the second hypothesis, namely that consistency of response to a core of items is related to profile stability, eight groups of items were selected for study.

The basis of this selection was largely a commonsense one. The writer simply examined the occupational, school subject, amusement, and activity items on the Strong blank and arbitrarily classified as many of these items as seemed feasible into broad occupational activities. While it is recognized that the method employed may introduce a sizeable bias, the alternative would appear to be to simply forget about the hypothesis since the enormous task of item-intercorrelation, underlying a more empirical method, has not been undertaken.

The classification scheme, together with the items selected, is shown in Table 9.

For each individual, the response "Like", "Indifferent", or "Dislike" to each item within a core was tabulated, and the modal response for each core determined. Two scores were then obtained--the modal score and the deviation score. The deviation score was computed by giving a weight of one for each response adjacent to the modal response and a weight of two for each response two spaces from the modal response. Again an example may be clarifying.

Subject #59

Area	Response			Modal Response	Modal Score	Deviation Score
	L	I	D			
Computational	~~1111~~ 111	~~1111~~	11	L	8	9

Table 9. Classification of items.

Core Group	Related Items
Computational:	16, 21, 22, 91, 101, 103, 105, 107, 115, 120, 148, 219, 223, 284, and 306.
Sales:	8, 11, 54, 77, 81, 90, 95, 99, 160, 197, 217, 245, and 287.
Social Service Work:	7, 26, 28, 41, 42, 48, 70, 85, 89, 100, 109, 133, 196, 197, 198, 206, and 209.
Biological Sciences:	29, 49, 67, 83, 93, 106, 108, 129, 130, 136, 162, and 192.
Aesthetic:	5, 39, 43, 50, 71, 84, 104, 163, 167, 193, 194, 286, and 305.
Manual Worker:	13, 19, 36, 46, 68, 73, 94, 98, 121, 132, 180, 186, 187, 188, 189, 190, and 282.
Supervisory:	7, 17, 22, 33, 35, 38, 41, 42, 58, 65, 70, 81, 88, 200, 206, 207, 285, and 311.
Clerical:	15, 16, 21, 31, 40, 57, 64, 74, 78, 79, 105, 112, 214, 215, 223, and 312.

In this case the mode is "L" and the number of "Likes" is eight; therefore the modal score is eight. For the deviation score, each "Indifferent" response is weighted one (total - five) and each "Dislike" response is weighted two (total - four). These are summed for a total of nine which is the deviation score.

This computation was done for each of the eight core groups of items. All eight scores were summed to yield the total modal and the total deviation score. These two "total" scores were used to test the second hypothesis.

The third hypothesis was that, using Strong's finding that certain items show a characteristic increase in liking and a decrease in disliking from age 15 to 25, a simple count of the number of likes or dislikes among these items might provide a meaningful index of profile stability.

For the purpose of testing this hypothesis, five parts of the Strong test were used, i.e., occupations, school subjects, amusements, activities, and possession of desirable traits.

A scoring key was prepared for each of these fields as follows:

Occupations - the number of dislikes indicated by each individual on items 1 through 100. ("Disliking" is more characteristic of 15 year olds than 25 year olds).

School Subjects - the number of dislikes indicated by each individual on items 101 through 136. ("Disliking" is more characteristic of 15 year olds that 25 year olds).

Amusements - the number of dislikes indicated by each individual on items 135 through 185. ("Disliking" is more characteristic of 15 year olds than 25 year olds).

Activities - the number of dislikes indicated by each individual on items 186 through 233 ("Disliking" is more characteristic of 15 year olds that 25 year olds).

Possession of Desirable Traits - on this scale there were 40 items. It was decided that six of these should be eliminated because they did not show the characteristic trend in liking or disliking noted above. The remaining 34 items were then scored by counting the "Likes" expressed by each individual on items 361 through 400 (except for items 374, 377, 389, 397, 398, and 400). These "Likes" were then subtracted from 34. This was necessary in order to arrive at a score which could be summed with the scores obtained

from the other parts. That is, on this part, the characteristic change from age 15 to 25 is from "Like" to "Indifferent" or "Dislike", whereas on the other parts the characteristic change is from "Dislike" to "Indifferent" or "Like". By subtracting the "Likes" from 34 (the number of items considered), a score was obtained which corresponded with those on the other parts and could be summed with them.

Scores thus obtained on these five parts were summed to yield the similarity score for each individual by which the third hypothesis was to be tested.

Table 10 has been prepared as a summary of the various proposed indices together with their computational method. The appendix includes copies of various forms used in making computations. These forms may provide additional clarification of the measurement procedures.

Table 10. Summary of predictive indices.

Index	Abbreviation	Computational Method
"Highs"	X_1	Comparison of five highest scores with each other.
"Lows"	X_2	Comparison of five lowest scores with each other.
"Highs vs Lows"	X_3	Comparison of each of the five highest scores with each of the five lowest scores.
Modal	X_4	Number of modal responses to a homogeneous group of items.
Deviation	X_5	Variability from modal response to a homogeneous group of items.
Similarity	X_6	Degree of similarity of response to typical responses of 15 year olds.

Criteria

Two criterion measures were employed. The first was a measure of stability between 1949 and 1951 testings, while the second covered the longer time period, 1949 to 1953. In both instances, the measure of stability employed was the rank order correlation (rho) between the two profiles (converted by the z' transformation. The validity of this measure has been established by Hoyt (17).

Method

Two methods of testing the hypotheses were used. First, scores on each of the predictive variables were correlated with both criterion measures. In addition, intercorrelations were computed as a matter of general interest and in order to compute multiple correlations between any combination of promising predictors.

Second, means and standard deviations for the three groups of students on each of the indices were computed, and tests of significance of the differences were made. These groups were defined by Stordahl, as noted above, on the basis of the rank order correlation between their 1949 and 1951 profiles, and were denoted as "High", "Average", and "Low" stability groups.

RESULTS AND DISCUSSION

Correlational Analysis

The major hypotheses were first tested by means of correlational analyses. As noted earlier, two criteria were employed: (a) interest stability over the two-year period, 1949 to 1951, measured by the z' equivalent of the rank order correlation for 44 occupations; and (b) interest stability over the

four-year period, 1949 to 1953, measured by an identical method. These two criteria will be referred to as Y_1 and Y_2 respectively in the succeeding discussion.

To review briefly the preceding section, the first hypothesis was that interest stability was related to the consistency of the individual's pattern of scores. Sub-hypothesis (a) stated that interest stability was related to the inter-relationships among the individual's five highest scores. The X_1 index was correlated with Y_1 and Y_2 as a test of this hypothesis. Sub-hypothesis (b) stated that interest stability was related to the inter-relationships among the individual's five lowest scores. The correlation of X_2 with Y_1 and Y_2 provided the test of this hypothesis. Sub-hypothesis (c) stated that interest stability was related to the inter-relationships among the individual's five highest and five lowest scores. For a test of this hypothesis, the correlation of X_3 with Y_1 and Y_2 was computed.

The second major hypothesis was that interest stability was related to item consistency. Two measures of item consistency, X_4 and X_5, were correlated with Y_1 and Y_2 to provide a statistical test of this hypothesis.

Finally, the third hypothesis stated that interest stability was related to the similarity in item responses to 25-year olds as opposed to 15-year olds. Statistically, a test of this hypothesis was provided by correlating X_6 with Y_1 and Y_2.

Table 11 presents the complete inter-correlation matrix for Y_1 and the six predictive indices.

The essential tests of the three major hypotheses with regard to Y_1 are found in the first line of this table. Other correlations are presented as a matter of general interest; they were used in computing multiple correlations.

Table 11. Intercorrelations between predictive indices and criterion measure using 176 subjects.

	X_1	X_2	X_3	X_4	X_5	X_6
Y_1	$.243^{**}$	$.202^{**}$	$-.328^{***}$	$.090^{ns}$	$-.152^{*}$	$-.030^{ns}$
X_1	--	$.174^{*}$	$-.413^{***}$	$.061^{ns}$	$-.142^{ns}$	$.030^{ns}$
X_2		--	$-.453^{***}$	$.160^{*}$	$-.069^{ns}$	$.154^{*}$
X_3			--	$-.100^{ns}$	$.038^{ns}$	$-.059^{ns}$
X_4				--	$-.043^{ns}$	$.150^{*}$
X_5					--	$.117^{ns}$
X_6						--

ns - Not significant
* Significant at .05 level.
** Significant at .01 level.
*** Significant at .001 level.

Clearly, from this table, the first hypothesis was substantiated. Although the correlations were not of a startling magnitude, they did indicate that there is less than 1 chance in 100 that the "true r" is zero. Furthermore, the correlation of X_1, X_2, and X_3 with Y_1 are all in the expected direction. It appears safe to conclude that there was statistical support for each of the sub-hypotheses as well as for the more general major hypothesis.

Regarding the second major hypothesis, the statistical test is less conclusive. One index, X_5, was correlated significantly with Y_1, but only at the .05 level of probability. In the case of X_4, the correlation was not significantly different from zero. As far as the direction of the correlation goes, X_4 was expected to correlate possitively while X_5 was expected to correlate negatively with the criterion. Although these expecta-

tions were fulfilled, the magnitude of the obtained r's is so unimpressive as to lead to the general conclusion that the second hypothesis was not substantiated.

There would appear to be little doubt regarding the third hypothesis. The correlation of -.030 between X_6 and Y_1 must be interpreted as a rejection of this hypothesis. The "similarity" scores obtained in this manner apparently have no more predictive value than did the similarity scores (Interest Maturity) utilized by Stordahl.

In Table 12 the three major hypotheses are tested against the Y_2 criterion. Only those intercorrelations involving indices which had some predictive power are included in this table.

Table 12. Partial intercorrelations between predictive indices and criterion measure using 116 subjects.

	X_1	X_2	X_3	X_4	X_5	X_6
Y_2	$.202^{**}$	$.243^{**}$	$-.373^{***}$	$.093^{ns}$	$-.100^{ns}$	$.036^{ns}$
X_1		$.230^{**}$	$-.525^{***}$	--	--	--
X_2			$-.472^{***}$	--	--	--
X_3				--	--	--

ns Not Significant.
* Significant at .05 level.
** Significant at .01 level.
*** Significant at .001 level.

Table 12 provides roughly the same answer as did Table 11. Again the X_1, X_2 and X_3 indices correlated significantly with the criterion and thus offered confirmation of the first hypothesis. These data appear more clearly to substantiate the earlier conclusion regarding the second hypothesis, namely

that it was not supported. Again X_6 was unrelated to the criterion leading to a rejection of the final hypothesis.

A perusal of Tables 11 and 12 leads to certain other observations. Indices X_1 and X_2 predicted interest stability about equally well, whereas X_3 had a higher correlation with the criteria. In addition, Y_2 was predicted equally well as Y_1 by X_1 and X_2. In the case of X_3, Y_2 was actually predicted better than Y_1. This may be significant, inasmuch as the Y_2 index was probably a more accurate measure of interest stability than was Y_1. That is, Y_2 represented the stability of profiles obtained during the senior year in high school and the senior year in college; Y_1 represented the stability obtained between the high school senior and the college sophomore profiles. It seems likely that the college senior profile represented the "final" interests of the subject better than the college sophomore profile did. If such was the case, Y_2 obviously would be a better criterion than Y_1.

Aside from the correlations with the criteria, the intercorrelations among the various "X" variables were of interest. In considering X_4, X_5, and X_6, these correlations were, with two exceptions not statistically different from zero. The highest, .160 between X_2 and X_4 was significant at the 5 per cent level; but in the practical sense, these variables, too, must be considered to be relatively independent.

On the other hand, X_1, X_2, and X_3 are significantly intercorrelated. The relationships tended to be slightly higher when the 116 subjects tested in 1953 were used than when the entire sample of 176 tested in 1951 were used. Although all of these intercorrelations were statistically significant, they were all relatively low. In fact, the highest was only - .525 (for X_1 and X_3 using 116 subjects). This suggested that a judicious combination of X_1, X_2, and X_3 might lead to a considerably improved prediction

of Y_1 and Y_2.

To investigate this possibility, multiple correlations between X_1, X_2, and X_3 on one hand, and Y_1 and Y_2 on the other, were computed. The resultant R's, .354 and .380 for Y_1 and Y_2 respectively, were not a significant improvement over the zero order r's obtained between X_3 and the criteria (-.328 and -.373). In other words, the X_1 and X_2 indices did not add any independent information, not available from X_3 about the two criteria.

In summary, the results of the correlational analyses provided confirmation for one hypothesis, but no support for the other two. Specifically, the hypothesis pertaining to consistency of occupational scores, and its three sub-hypotheses, were confirmed by the data. The hypotheses pertaining to consistency of item response and similarity of item response were not substantiated.

None of the variables found to have statistically significant correlations with the criteria could be said to have practical significance. That is, the correlations were all of such a magnitude that little could be inferred about the probable interest stability for any given individual from consideration of his X_1, X_2, or X_3 scores. Furthermore, the optimal combination of these variables did not significantly improve their predictive capacity.

Mean Scores

In the interest of clarification, it is worthwhile to consider a second method of testing the three hypotheses. In this method the sample was subdivided into three stability groups - "High", "Average", and "Low"--using Stordahl's criterion. Mean scores on each index for these three groups were

then computed, and a *t* test run between the "High" and "Low" groups.

For the 176 students who were retested as college sophomores there were 60, 61, and 55 in the "High", "Average", and "Low" group respectively. Of these, there were 39, 39, and 38, respectively, who were retested as college seniors. Computations were done for both sets of these three groups on each index.

In Table 13, the results of these computations for X_1 are presented.

Table 13. Scores on the X_1 index for groups with varying degrees of interest stability over two and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	19.35	9.33
	4-year	39	21.00	9.27
Average	2-year	61	16.38	9.41
	4-year	39	17.20	9.51
Low	2-year	55	15.15	9.41
	4-year	38	16.29	9.28

The mean scores obtained by the High and Low groups are significantly different ($P<.02$) for both the two-year and the four-year intervals. As in the correlational analysis, sub-hypothesis (a) was substantiated.

Three incidental observations from Table 13 seem worthy of note. First of all, the variability in X_1 scores among the groups and across time-intervals was highly uniform. Secondly, the Average group obtained a mean score in between that of the High and Low groups, but appeared much closer to the latter than to the former. Thirdly, mean scores for the four-year interval groups were uniformly higher than for the two-year groups. These observations will be commented on at greater length later on in this report.

Table 14 provides similar data regarding the X_2 index.

Table 14. Scores on the X_2 index for groups with varying degrees of interest stability over two and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	16.45	8.80
	4-year	39	17.95	8.59
Average	2-year	61	13.30	8.55
	4-year	39	12.77	8.03
Low	2-year	55	13.42	9.86
	4-year	38	13.92	10.10

Here, as in the case of the X_1 index, the mean scores obtained by the High and Low groups are significantly different ($P < .05$) for both the two-year and the four-year intervals. While the level of confidence is not as high as in the case of the X_1 index, it is adequate to offer statistical support for sub-hypothesis (b).

Once again, several incidental observations seem worthy of note. First of all, the variability in X_2 scores was fairly uniform across time intervals and among groups. The Low group was insignificantly more variable than the other two groups for both time intervals. The mean score of the Average group was below that of the other groups, although again it was quite similar to that of the Lows. Finally, the mean scores for the four-year interval tended to be slightly higher than for the two-year interval.

Table 15 provides the same data for the X_3 index.

Mean scores obtained by the High and Low groups are significantly different ($P < .01$) for both the two-year and the four-year intervals. This, as in the case of the correlational analysis, substantiated sub-hypothesis (c).

Table 15. Scores on the X_3 index for groups with varying degrees of interest stability over two and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	-27.43	19.13
	4-year	39	-31.64	16.35
Average	2-year	61	-17.56	16.33
	4-year	39	-17.41	15.49
Low	2-year	55	-16.77	14.94
	4-year	38	-16.21	16.37

The magnitude of the difference between mean scores makes this finding more impressive than the findings for X_1 and X_2. While variability of the X_3 scores among groups for the four-year interval was uniform, the Highs were significantly more variable than the Lows for the two-year interval. This necessitated the substitution of the Behrens-Fisher *d* test (19) for the more usual *t* test of the significance of the difference in mean scores. The mean score of the Average groups fell in between those of the High and Low groups but again are much closer to the latter than to the former. Finally mean scores for the four-year interval groups tended to be slightly higher than for the two-year groups. This was a function entirely of the High group, however.

It will be noticed in the observations concerning all three of these indices that the mean scores of the Average group were closer to the Low group than to the High group. This suggests that either a clear cut distinction cannot be made between those with Average stability and those with Low stability or that the predictive indices bear a curvilinear, rather than a rectilinear, relationship to the criteria. At any rate, the predictive

indices did not appear to be equally valid throughout the entire stability range.

With one exception, the variability of the scores among groups and across time-intervals was fairly uniform. In the case of the exception (the Highs were more variable than the Lows on X_3 for the two-year interval), the differential variability washed out for the four-year interval.

Mean differences between the High and Low groups varied from about one-third of a standard deviation to almost one full standard deviation. These differences were most marked for the X_3 index, and least impressive in the case of X_2.

Mean scores for the four-year interval appeared to be slightly higher than for the two-year interval. This seemed to be due primarily to trends within the High's group. Here again no definite pattern was apparent in the case of the Average or Low groups. It would seem, on the basis of these observations, that these scales were sufficiently sensitive to distinguish between High groups and Low groups. On the other hand, they were not sensitive enough to discriminate between Average and Low groups. Whether this finding reflects a defect in the scales or in the definition of the criterion groups cannot be answered at this time.

In brief summary of these results, the first hypothesis, and its three sub-hypotheses, were confirmed statistically. The X_3 index provided a better differentiation between High and Low groups than did either the X_1 or X_2 indices. None of the scales adequately differentiated the Average and Low groups.

The second hypothesis was tested in the same manner, using indices X_4 and X_5. Tables 16 and 17 give the results of these computations.

Table 16. Scores on the X_4 index for groups with varying degrees of interest stability over two-and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	73.42	9.30
	4-year	39	73.28	8.10
Average	2-year	61	67.80	9.43
	4-year	39	67.13	11.02
Low	2-year	55	70.09	8.93
	4-year	38	72.00	9.19

Table 17. Scores on the X_5 index for groups with varying degrees of interest stability over two- and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	57.13	14.31
	4-year	39	57.87	13.15
Average	2-year	61	61.57	12.07
	4-year	39	61.95	13.33
Low	2-year	55	62.15	14.37
	4-year	38	59.52	14.55

On both indices, the mean scores obtained by the High and the Low groups are not significantly different for either the two-year or the four-year interval. As in the correlational analysis the second hypothesis was not substantiated. Again an academic question regarding the significance of X_5 was raised. As noted in Table 17, for the two-year interval the mean scores do array themselves properly from the High group to the Low group. The mean difference approached the 5 per cent level of statistical significance. However, this slight trend was obliterated in the four-year group.

Thus the general conclusion was that the second hypothesis could not be sustained.

Table 18 provides similar data regarding the X_6 index used in testing the final hypothesis.

Table 18. Scores on the X_6 index for groups with varying degrees of interest stability over two- and four-year periods.

Stability Group	Interval	N	Mean	Standard Deviation
High	2-year	60	92.03	37.05
	4-year	39	94.00	38.58
Average	2-year	61	88.89	28.05
	4-year	39	91.08	25.58
Low	2-year	55	93.47	35.87
	4-year	38	94.45	38.76

Mean scores obtained by the High and Low groups were not significantly different for either the two-year or the four-year intervals. Neither did the Average group score in between the two extreme groups. Thus, as in the case of the correlational analysis, the third hypothesis was not supported.

Inasmuch as these results confirmed those obtained by correlational analysis, it was concluded that the second and third hypotheses were not tenable. The X_4, X_5, and X_6 indices showed so little promise as predictors of interest stability that they were omitted from further analysis.

In brief summary, the results of correlational analyses and comparison of the means of High and Low groups confirmed the first hypothesis but did not substantiate the other two. Specifically the hypothesis pertaining to consistency of occupational scores was confirmed by the data. The hypotheses pertaining to consistency of item response and similarity of item response

were not substantiated. In view of these results, indices X_4, X_5, and X_6 were eliminated from further analysis.

Side Issues

Although not germane to the major purpose of this investigation, namely the testing of the three major hypotheses, some further analysis of the X_1, X_2, and X_3 indices and the criterion groups seemed appropriate in order to shed some light on certain side issues. Four such issues were investigated: (a) the relationship of these indices to college persistence; (b) the definition of the four-year interval criterion groups; (c) the practical meaning of the results reported above; and (d) changes in these indices over time.

College Persistence. No known theory of interests has postulated college persistence to be a function of interest stability. Yet it occurred to the writer that in some cases, discouragement over not being able to "find" one's interests might well lead to withdrawal from college for a time. Some of the data at hand could provide for a tentative exploration of this question.

It will be recalled that Hoyt tested only 116 of the 176 students in Stordahl's group. The group of 60 not tested was made up of 40 drop-outs and 20 who could not be persuaded to take the Strong in 1953. It was impossible, however, to identify which were drop-outs and which were "non-cooperators", so the entire group of 60 was labeled a "drop-out" group. The remaining 116 were labeled the "persistent" group.

Means and standard deviations were computed for both groups on indices X_1, X_2, and X_3. Table 19 shows the results of these computations.

For all three indices, the persisters obtained higher mean scores than the drop-outs. However the difference in means was significant only in the

Table 19. Scores on three indices of interest stability for college "persisters" and college "drop-outs".

Index		"Persisters" N = 116	"Drop-outs" N = 60
X_1	Mean	18.18	14.73
	s.d.	9.50	9.51
X_2	Mean	14.89	13.48
	s.d.	9.13	9.24
X_3	Mean	-21.80	-20.70
	s.d.	17.42	17.94

case of the X_1 index ($P < .02$). The results tended to confirm the hypothesis, but further studies with better defined criterion groups are needed before definite conclusions can be reached.

Definition of Stability Groups. It will be recalled that the three stability groups over a four-year period were defined in terms of two-year interval stability scores. That is, the Highs were simply those students from the two-year High group who had been retested in 1953.

The adequacy of this definition depends upon how well two- and four-year stability figures agree with each other. On the assumption that there would be enough disagreement to affect the mean scores of the groups, new High, Average, and Low groups were defined on the basis of the rank order correlation between 1949 and 1953 profiles.

For this purpose, Hoyt's (17) findings were used. It will be recalled that when the rho between profiles was .75 or higher there was no major change in the interpretation of the profile; but when the profile rho was .65 or less, major changes were apparent. On the basis of this finding, those with a rho of .75 or higher were included in the High group; those

with a rho of .65 or less were included in the Low group; and profiles with rhos between these figures were included in the Average group.

Means and standard deviations were computed for each of these stability groups on indices X_1, X_2, and X_3. Comparisons were made with the results obtained using Stordahl's defined groups.

In Table 20 the results of these computations for X_1 are presented.

Table 20. Comparison of scores on the X_1 index obtained by interest stability groups defined by two different methods.

Stability Group	Defined by	N	Mean	Standard Deviation
High	Stordahl	39	21.00	9.27
	Smith	42	19.88	8.53
Average	Stordahl	39	17.21	9.51
	Smith	36	20.14	8.82
Low	Stordahl	38	16.29	9.27
	Smith	38	14.71	10.21

By inspection, the two methods of defining stability groups led to the same general results. Curiously, the revised definition resulted in an insignificantly higher mean score for the Average group than for the High group. With Stordahl's definition, Average and Low groups tended to score alike. Aside from this, the results appeared quite similar for both sets of criterion groups. It might be pointed out, however, that in the differences of the means of the Highs and Lows, Stordahl's figures were significant at the .02 level whereas this investigator's were significant at the .01 level.

Table 21 presents similar data for the X_2 index.

Once more there were no significant differences between the means of the groups as defined by Stordahl and the means of the groups defined by Smith.

Table 21. Comparison of scores on the X_2 index obtained by interest stability groups defined by two different methods.

Stability Group	Defined by	N	Mean	Standard Deviation
High	Stordahl	39	17.95	8.59
	Smith	42	18.22	9.02
Average	Stordahl	39	13.30	8.55
	Smith	36	13.39	7.89
Low	Stordahl	38	13.42	9.86
	Smith	38	13.03	9.71

The differentiation does appear to be a little sharper for the new criterion groups than for Stordahl's. This impression was confirmed by a difference in the significance levels--.05 for Stordahl's groups and .01 for Smith's.

Table 22 presents the data which were obtained for the X_3 index.

Table 22. Comparison of scores on the X_3 index obtained by interest stability groups defined by two different methods.

Stability Group	Defined by	N	Mean	Standard Deviation
High	Stordahl	39	-31.64	16.35
	Smith	42	-29.10	16.78
Average	Stordahl	39	-17.41	15.49
	Smith	36	-21.56	15.41
Low	Stordahl	38	-16.21	16.37
	Smith	38	-15.08	16.67

Here, as in the case of the other indices, the differences between the means of the groups as defined by Stordahl and Smith were insignificant. The differences in the means of the Highs and Lows for both sets were significant well beyond the .01 level. The Average group was somewhat better behaved for

Smith than for Stordahl. By Smith's definition, the mean for this group almost splits the difference between the Highs and the Lows. As was pointed out earlier, Stordahl's Average group scored a good deal like his Lows.

All in all, there were only very slight differences in the mean scores of the two sets of criterion groups. The more exacting and current definition used by this investigator appeared to have a slight advantage over Stordahl's in terms of the sharpness with which the groups could be distinguished on the three criteria.

Practicality. Throughout this section, references have been made to the fact that while the results obtained with X_1, X_2, and X_3 were statistically significant, there did not appear to be a correspondingly high degree of practical significance. This point is herein examined more fully.

For this purpose, computations of the per cent of overlap have been made. Two sets of over-lap figures were obtained--(1) the per cent of the High group who reached or fell below the mean of the Low group; and (2) the per cent of the Low group who reached or exceeded the mean of the High group. Separate computations were made for Stordahl's two- and four-year groups and Smith's four-year group.

Table 23 presents the results obtained for the High groups on the three indices of stability. Several generalizations are apparent.

1. Most importantly, there was a good deal of overlap present on each index regardless of how the High group is defined. The most impressive figure, 17.9 per cent, still indicates that nearly 1 of 5 in the High group scores at or below the mean of the Low group. This margin of error is too high to permit serious use of these indices with individual students.

2. Of the three indices, X_2 consistently misclassified a greater number of students than did either X_1 or X_3. As in other comparisons, X_3 was

the most valid index.

3. There were slightly fewer misclassifications over the four-year interval than over the two-year interval. This trend was more evident for X_1 and X_2 than for X_3.

Table 23. Per cent of the high stability group who reach or fall below the mean score of the low stability group on three indices of interest stability.

Index	Interval	Defined by	Per cent
X_1	2-year	Stordahl	28.3
	4-year	Stordahl	25.6
	4-year	Smith	21.4
X_2	2-year	Stordahl	43.3
	4-year	Stordahl	36.4
	4-year	Smith	35.7
X_3	2-year	Stordahl	23.3
	4-year	Stordahl	17.9
	4-year	Smith	23.8

Table 24 presents similar data for the Low stability group.

With one exception, the generalizations from Table 23 held true here. In this instance, X_2 appeared slightly more valid than X_1. Again X_3 made the fewest misclassifications.

Changes with Time in the Predictive Indices. One final analysis of the stability scales was made. This was based on the fact that interests tend to become more stable over time. It seemed to be a logical hypothesis that indices for predicting stability should become higher (predict more stability) on successive testings. To test this hypothesis a comparison was made of the mean scores obtained on X_1, X_2, and X_3 at two different testing times--1949 and

Table 24. Per cent of the low stability group who reach or exceed the mean score of the high stability group on three indices of interest stability.

Index	Interval	Defined by	Per cent
X_1	2-year	Stordahl	36.4
	4-year	Stordahl	31.6
	4-year	Smith	34.2
X_2	2-year	Stordahl	35.9
	4-year	Stordahl	31.6
	4-year	Smith	28.9
X_3	2-year	Stordahl	27.3
	4-year	Stordahl	21.1
	4-year	Smith	23.7

1951. Means and standard deviations were computed for each of the stability groups, and for the total group, at each testing time for the three indices of stability.

Table 25 shows the results obtained for the X_1 index.

Table 25. Comparison of scores on the X_1 index of interest stability obtained by three stability groups at two different testing times.

Stability Group	N	Testing Period			
		1949		1951	
		Mean	s.d.	Mean	s.d.
High	60	19.35	9.33	20.52	7.11
Average	61	16.38	9.41	16.97	8.45
Low	55	15.15	9.41	14.04	8.29
Total	176	17.01	9.58	17.26	8.34

There were no significant differences between 1949 and 1951 mean scores for any of the groups or for the total group. With the exception of the

Low stability group, trends in the direction predicted by the hypothesis were apparent; but the differences were too slight to be of statistical significance.

Table 26 presents similar data for the X_2 index.

Table 26. Comparison of scores on the X_2 index of interest stability obtained by three stability groups at two different testing times.

Stability Group	N	Testing Period			
		1949		1951	
		Mean	s.d.	Mean	s.d.
High	60	16.45	8.80	18.13	7.76
Average	61	13.30	8.55	17.03	9.09
Low	55	13.42	9.86	14.71	9.49
Total	176	14.41	9.13	16.68	8.85

In this instance all of the group means did change in the direction expected. The differences were significant only in the case of the Average group and for the total group. Thus the data for X_2 provided some support to the hypothesis.

In Table 27 the data obtained for the X_3 index are presented.

Table 27. Comparison of scores on the X_3 index of interest stability by three stability groups at two different testing times.

Stability Group	N	Testing Period			
		1949		1951	
		Mean	s.d.	Mean	s.d.
High	60	-27.43	19.13	-36.35	16.12
Average	61	-17.56	16.33	-28.95	14.53
Low	55	-16.77	14.94	-24.56	16.08
Total	176	-21.71	17.53	-30.10	16.23

For all three stability groups, and for the total, the differences between 1949 and 1951 means were significant well beyond the .01 level. The most valid index, X_3, thus offered the most convincing support for the hypothesis.

Taking the results from all three scales into account, it seems fair to generalize by saying the expected increase in the predictive indices did occur, although it was slight except in the case of X_3.

SUMMARY, CONCLUSIONS AND IMPLICATIONS

Summary

This study investigated three major hypotheses concerning the prediction of the stability of vocational interests as measured by the Strong Vocational Interest Test. These hypotheses were:

1. A stable Strong Vocational Interest Profile (one that is reliable over time) will show a higher order of consistency among its scales than an unstable one. Three specific sub-hypotheses were derived from this.

   a. The relationships among the five highest occupational scores will be of a higher order for stable profiles than for unstable ones.

   b. The relationships among the five lowest occupational scores will be of a higher order for stable profiles than for unstable ones.

   c. The relationships among the five highest scores, on one hand, and the five lowest, on the other, will be of a greater negative magnitude for stable profiles than for unstable ones.

2. Consistency of response to a core of similar items will be related to profile stability.

3. Similarity in item response to 25-year olds as opposed to 15-year

olds will be more characteristic of stable profiles than of unstable ones.

The sample used for this study was 176 of the 181 students used by Stordahl (33). These students were originally tested on the Strong in 1949 as high school seniors. Stordahl retested them in 1951 as college sophomores. For a portion of the experiment, 116 of these students who were tested for a third time in 1953 by Hoyt composed the sample.

Two criteria were employed. The first was the z' equivalent of the rank order correlation between 1949 and 1951 profiles. The second was the same measure applied to 1949 and 1953 profiles. Earlier work had established the z' index as a valid indicator of interest stability.

Six predictive indices were devised on the basis of the hypotheses. Statistical tests of these hypotheses were made by two methods:

1. Correlational analysis.
2. Analysis of mean score differences.

In the correlational analysis, each of the six predictive indices was correlated with both criteria. For the second method of testing the hypotheses (analysis of mean score differences), groups of "High", "Average", and "Low" stability were defined on the basis of z' scores. Means and standard deviations for each of the three groups on each of the six indices were computed, and t tests of the significance of the difference between the means of the "High" and "Low" groups were run.

Conclusions

Within the limits of the sample used and with the reservations noted under "Implications", the following conclusions seem warranted:

1. The first hypothesis was supported. A statistically significant

relationship was established between each of the three predictive indices and both criteria. This conclusion held true regardless of the method used to test the hypothesis.

2. Despite the general confirmation of the first hypothesis, the magnitude of the relationship between the predictive indices and the criteria was unimpressive. Coefficients of correlation varied from .20 to .37. Overlap varied from 18 per cent to 43 per cent.

3. Of the indices, the one designed to test sub-hypothesis (c) was consistently more closely related to the criteria than either of the other two.

4. Optimally combining the three predictive variables through the multiple correlation method did not substantially increase their relationship to the criteria.

5. Stability over a 4-year period was predicted slightly more accurately than stability over a 2-year period.

6. The second hypothesis was largely unsubstantiated by the data. In general, the trend of the results was in line with the hypothesis, but statistical significance (and practical significance) was lacking.

7. No support was found for the third hypothesis. Similarity in response to 25-year olds as opposed to 15-year olds apparently had no more significance for interest stability in this case than it did in the case of Strong's Interest Maturity scale.

In addition to the seven conclusions listed above, certain conclusions regarding the three predictive indices used to test the first hypothesis were apparent.

8. These indices were insignificantly more closely related to interest

stability over a 4-year period when various stability groups were defined by an exacting criterion than when they were defined by the degree of stability over a 2-year period.

9. There was a tendency for college "persisters" to score higher on these indices than for college "drop-outs". Statistical significance was apparent for only one of the three indices, however. This finding is very tentative for lack of a pure "drop-out" group.

10. Scores on the three predictive indices were higher in 1951 than in 1949. This was a statistically reliable trend for only the second and third index. Inasmuch as interests do tend toward stability with increasing age, this finding offered additional indirect support to the first hypothesis.

Implications

In general, the results of this investigation were gratifying. For the first time, indices which will predict the stability of Strong patterns significantly better than chance have been derived.

On the other hand, much work remains to be done before research can provide counselors with a measure of probable interest stability which will have meaning in individual cases. The difference between statistical and practical significance cannot be overlooked.

What directions might future research take to close the gap between available measures and counseling requirements? Several suggestions might be made on the basis of this investigation.

1. The indices used in testing the first hypothesis might well be improved if the complete table of intercorrelations for Strong's scales were available. Strong's table omitted data on the revised psychologist key, the

relationship was established between each of the three predictive indices and both criteria. This conclusion held true regardless of the method used to test the hypothesis.

2. Despite the general confirmation of the first hypothesis, the magnitude of the relationship between the predictive indices and the criteria was unimpressive. Coefficients of correlation varied from .20 to .37. Overlap varied from 18 per cent to 43 per cent.

3. Of the indices, the one designed to test sub-hypothesis (c) was consistently more closely related to the criteria than either of the other two.

4. Optimally combining the three predictive variables through the multiple correlation method did not substantially increase their relationship to the criteria.

5. Stability over a 4-year period was predicted slightly more accurately than stability over a 2-year period.

6. The second hypothesis was largely unsubstantiated by the data. In general, the trend of the results was in line with the hypothesis, but statistical significance (and practical significance) was lacking.

7. No support was found for the third hypothesis. Similarity in response to 25-year olds as opposed to 15-year olds apparently had no more significance for interest stability in this case than it did in the case of Strong's Interest Maturity scale.

In addition to the seven conclusions listed above, certain conclusions regarding the three predictive indices used to test the first hypothesis were apparent.

8. These indices were insignificantly more closely related to interest

4. Again on a commonsense basis, it seems logical to suppose that item familiarity would be related to stability. A large number of Strong items are outside the range of the average college freshman's experience. Many are probably not even in his vocabulary. It would seem to be a simple, yet potentially productive, matter to construct a scale of item familiarity and relate the results to stability scores.

5. Insofar as interests are a function of experience, then the breadth of experience should be correlated with interest stability. It might be hypothesized that experience functions as a source of information to the individual about himself. By experiences, he learns that he likes this but dislikes that. Without experience, he has only tentative opinions easily modified as experiences are made available.

It would seem appropriate to test this type of reasoning by relating some measure of "breadth and/or intensity of experiences" to interest stability.

6. The above suggestions involve either statistical refinements or commonsense deductions. It would be much more desirable if suggestions could be made on the basis of sound knowledge of the determinants of interests. Such knowledge is unfortunately absent at the present time. This is probably a joint function of the elusive character of interests, the immensity of the empirical research task, and the lack of a systematic, coherent theory of interests. Whatever the causes, the lack of this basic knowledge will probably preclude the possibility of building a highly valid predictor of interest stability.

A noteworthy suggestion of Darley and Hagenah (6) in this connection bears repetition here. An intensive clinical study of individuals whose interests have been shown to be highly stable or highly unstable may prove

productive of hypotheses regarding interest stability. Such a study might investigate, at least tentatively, numerous developmental factors--family background, hobbies, work and school experiences, friendship patterns--as well as outcomes of these factors--personality needs, general and specific attitudes and beliefs, interpersonal patterns. A wealth of hypotheses would almost certainly result from an investigation of this type.

In passing, it should be noted that the long-range study of Leona Tyler (52) can be expected to increase the basic information about the development of interests immeasurably, although it will be a number of years before these findings will be available.

These suggestions obviously do not exhaust the possibilities. They represent the biases of the investigator, and are limited by his lack of knowledge and blind spots. If they serve to stimulate a single research in the general area of interest development or interest stability, this investigation will have fulfilled an important function.

ACKNOWLEDGMENTS

The author wishes to express his sincere but inadequate gratitude to Dr. Donald P. Hoyt without whose help and guidance at each stage of the project, this research and thesis could not have been completed.

The writer is also grateful to Dr. Seymour Levy who proposed the only fruitful hypothesis of this study.

Dr. Kalmer E. Stordahl made available the raw data on his 1949-1951 testings which greatly facilitated the project. Dr. Henry Tucker furnished statistical assistance which was also of great help.

Finally, to the writer's wife, Miriam, heartfelt thanks are due for her understanding and help during this undertaking.

To all of these people the author extends his sincere and deep appreciation for their contributions to this paper.

REFERENCES

(1) Barnett, G. J., Lawrence Stewart and Donald E. Super.
Level of occupational interest: dead weight or dynamism? Educ. Psychol. Measmt., 1953, 13, 193-208.

(2) Bordin, E. S. and Earl H. Wilson.
Change in interest as a function of shift in curricular orientation. Educ. Psychol. Measmt., 1953, 13, 297-307.

(3) Canning, L., Katherine Van F. Taylor and Harold D. Carter.
Permanence of vocational interests in high school boys. J. educ. Psychol., 1941, 32, 481-494.

(4) Carter, H. D.
Vocational interests and job orientation--a ten year review. Appl. Psychol. Mono., 1944, No. 2.

(5) Christensen, T. E.
Some observations with respect to the Kuder Preference Record. J. educ. Res., 1946, 40, 96-107.

(6) Darley, J. G. and Theda Hagenah.
Vocational interest measurement, theory and practice. Minneapolis: Univ. of Minn. Press, 1955.

(7) DiMichael, S. G. and D. H. Dabelstein.
Work satisfaction and work efficienty of vocational counselors as related to measured interests. Amer. Psychol., 1947, 2, 342-343.

(8) DiMichael, S. G.
Interest inventory results during the counseling interview. Occup., 1951, 30, 93-97.

(9) Dressel, Paul L.
Interests--stable or unstable. J. educ. Res., 1954, 48, 95-102.

(10) Dyer, D. T.
Relation between vocational interests of men in college and their subsequent occupational histories for ten years. J. appl. Psychol, 1939, 23, 280-288.

(11) Finch, F. H.
The permanence ofvocational interests. Psychol. Bull., 1935, 32, 682.

(12) Fowler, F. M.
Interest measurement-questions and answers. School Life, 1945, 28, 25-29.

(13) Fox, W. H.
The stability of measured interests. J. educ. Res., 1947, 41, 305-310.

(14) Gordon, H. C. and W. W. Herkness.
Do vocational interest questionnaires yield consistent results? Occup., 1942, 20, 424-429.

(15) Green, R. F.
Does a selection situation induce testees to bias their answers on interest and temperament tests? Educ. Psychol. Measmt., 1951, 11, 503-515.

(16) Herzberg, F., Arthur Banton and Betty Jo Steiner.
Studies of the stability of the Kuder Preference Record. Educ. Psychol. Measmt., 1954, 14, 90-100.

(17) Hoyt, D. P.
Measurement and prediction of the permanence of interests. Proceedings of the conference on interest measurement, Univ. of Minn. Press, 1954.

(18) Jackson, J.
A note on the crystallization of vocational interests. J. soc. Psychol., 1947, 26, 125-130.

(19) Johnson, Palmer O.
Statistical methods in research. New York: Prentice Hall Inc., 1949.

(20) Klugman, S. F.
Permanence of clerical interest in relation to age and various abilities. J. soc. Psychol., 1945, 21, 115-120.

(21) Kapp, T. and L. Tussing.
The vocational choices of high school students as related to scores on vocational interest inventories. Occup., 1947, 25, 334-339.

(22) Longstaff, H. P.
Fakability of the Strong Interest Blank and the Kuder Preference Record. J. appl. Psychol., 1948, 32, 360-369.

(23) Mallinson, G. G. and W. M. Crumrine.
An investigation of the stability of interests of high school students. J. educ. Res., 1952, 45, 369-383.

(24) McCoy, R. A.
Stability and change of measured vocational interests of high school students. Unpub. Thesis (Ed. D.), 1954, Univer. of Missouri.

(25) Powers, Mabel K.
A longitudinal study of vocational interests during the depression years. Ph. D. Thesis, 1954, Univer. of Minnesota.

(26) Rectenwald, E.
Attitudes toward occupations before and after vocational information. Occup., 1946, 24, 220-223.

(27) Reid, J. W.
Stability of measured interests in young adults. J. educ. Res., 1951, 45, 307-312.

(28) Rosenberg, N.
Stability and maturation of Kuder interest patterns during high school. Educ. Psychol. Measmt., 1953, 13, 449-458.

(29) Seagoe, May Y.
Permanence of interest in teaching. J. Educ. Res., 1945, 38, 678-684.

(30) Shoemaker, W. L.
Rejection of measured vocational interest areas by high school students. Unpub. Thesis (Ed. D.), 1955, Univer. of Missouri.

(31) Silvey, H. M.
Changes in test scores after two years in college. Educ. Psychol. Measmt., 1951, 11, 494-502.

(32) Steinmetz, H. L.
Measuring ability to fake occupational interest. J. appl. Psychol., 1932, 16, 123-130.

(33) Stordahl, K. E.
Permanence of interests and interest maturity. J. appl. Psychol., 1954, 38, 339-341.

(34) Stordahl, K. E.
Permanence of Strong vocational interest blank scores. J. appl. Psychol., 1954, 38, 423-427.

(35) Strong, E. K., Jr.
Permanence of vocational interests. J. educ. Psychol., 1934, 25, 336-344.

(36) Strong, E. K., Jr.
Predictive values of the vocational interest test. J. educ. Psychol., 1935, 26, 331-349.

(37) Strong, E. K., Jr.
Vocational interests of men and women. Stanford: Stanford Univ. Press, 1943.

(38) Strong, E. K., Jr.
The role of interests in guidance. Occup., 1949, 27, 517-522.

(39) Strong, E. K., Jr.
Permanence of interest scores over 22 years. J. Appl. Psychol., 1951, 35, 89-91.

(40) Strong, E. K., Jr.
Interest scores while in college and college occupations twenty years later. Educ. Psychol. Measmt., 1951, 11, 335-348.

(41) Strong, E. K., Jr.
Amount of change in occupational choice of college freshmen. Educ. Psychol. Measmt., 1952, 12, 677-691.

(42) Strong, E. K., Jr.
19 year follow-up of engineer interests. J. appl. Psychol., 1952, 36, 65-74.

(43) Strong, E. K., Jr.
Vocational interests 18 years after college. Minneapolis: Univer. of Minn. Press, 1955.

(44) Super, D. E.
Vocational interests and vocational choice--present knowledge and future research in their relationships. Educ. Psychol. Measmt. 1947, 7, 375-383.

(45) Super, D. E.
Appraising vocational fitness. New York: Harper and Brothers, 1949.

(46) Taylor, Katherine V. F. and H. D. Carter.
Retest consistency of vocational interest patterns. J. consult. Psychol., 1942, 6, 95-101.

(47) Taylor, Katherine V. F.
The reliability and permanence of vocational interests of adolescents. J. exp. Psychol., 1942, 11, 81-87.

(48) Thorndike, E. L.
The interests of adults. I. The permanence of interests. J. educ. Psychol., 1935, 26, 401-410.

(49) Traphagen, A. L.
Interest patterns and retention and rejection of vocational choice. J. appl. Psychol. 1952, 36, 182-185.

(50) Traxler, A. E. and W. E. McCall.
Some data on the Kuder Preference Record. Educ. Psychol. Measmt., 1941, 1, 253-268.

(51) Trinkhaus, W. K.
The permanence of vocational interests of college freshmen. Educ. Psychol. Measmt., 1954, 14, 641-647.

(52) Tyler, Leona E.
The development of interests. Proceedings of the conference on interest measurement, Univer. of Minnesota, 1955.

(53) Van Dusen, A. C.
Permanence of vocational interests. J. educ. Psychol., 1940, 31, 401-424.

APPENDIX

STATE COLLEGE OF AGRICULTURE AND APPLIED SCIENCE
LIBRARY
MANHATTAN
KANSAS

Form 1. Numerical values for intercorrelations of occupations.

Form 1. Numerical values for intercorrelations of occupations.

	Artist	Architect	Physician	Osteopath	Dentist	Mathematician	Physicist	Engineer	Chemist	Production Manager	Farmer	Aviator	Carpenter	Printer	Math., Phys. Sci. Teacher	Policeman	Forest Service Man	Y.M.C.A. Phys. Director	Personnel Director	Y.M.C.A. Secretary	Soc. Sci. H. S. Teacher	City School Supt.	Minister	Musician	C.P.A.	Accountant	Office Man	Purchasing Agent	Banker	Sales Manager	Real Estate Salesman	Life Insurance Salesman	Advertising Man	Lawyer	Author-Journalist	President - Mfg. Concern
Artist		4	3	1	3	3	2	1	2	-2	0	0	0	0	0	-1	0	-1	-2	-1	-1	0	1	2	0	-3	-3	-3	-3	-2	-1	-1	2	1	4	0
Architect	4		3	2	3	3	3	2	3	0	1	1	1	1	0	0	1	-1	-2	-2	-2	-1	0	2	0	-2	-3	-2	-3	-3	-2	-2	0	0	2	0
Physician	3	3		3	4	3	3	2	3	0	1	1	1	1	0	0	1	0	-2	-2	-2	-1	0	2	0	-3	-3	-2	-3	-3	-2	-2	0	0	2	0
Osteopath	1	2	3		3	1		1	2	0	1		1	2	2	2	2	2	0	0	0	0	1	2	-1	-1	-1	-2	-2	-2	-2	-1	-1	0	0	-1
Dentist	3	3	4	3		2	3	2	3	0	2	2	2	2	2	1	2	0	-2	-2	-2	-1	0	2	-1	-2	-2	-2	-2	-3	-3	-2	-1	0	1	0
Mathematician	3	3	3	1	2		4	3	4	0	1	1	1	1	1	0	1	-1	-2	-2	-2	0	0	1	0	-1	-2	-2	-2	-3	-3	-3	-1	0	1	-1
Physicist	2	3	3		3	4		4	4	1	2		2	2	2	0	2	-1	-2	-2	-3	-2	0	1	0	-1	-2	-1	-3	-3	-3	-4	-1	-1	1	0
Engineer	1	2	2	1	2	3	4		4	3	3	3	3	2	2	1	2	-1	-1	-3	-3	-2	-1	0	0	0	-1	0	-2	-2	-3	-3	-2	-2	0	0
Chemist	2	3	3	2	3	4	4	4		2	2	3	3	2	2	1	2	0	-1	-2	-2	-2	0	1	0	0	-2	-1	-3	-3	-4	-4	-2	-1	0	0
Production Manager	-2	0	0	0	0	0	1	3	2		2	3	2	1	2	2	2	0	0	-1	-2	-2	-2	-2	0	2	1	2	0	0	-1	-1	-3	-3	-3	1
Farmer	0	1	1	1	2	1	2	3	2	2		3	4	3	3	2	4	0	-2	-1	-1	-2	-1	0	-2	0	0	0	0	-2	-2	-3	-3	-3	-1	1
Aviator	0	1	1		2	1		3	3	3	3		3	3	3	3	3	0	0	-2	-2	-2	-1	0	-2	0	0	0	-1	-2	-2	-3	-2	-2	-1	0
Carpenter	0	1	1	1	2	1	2	3	3	2	4	3		3	3	3	3	0	-1	-1	-1	-2	0	0	-2	0	0	0	-1	-2	-2	-3	-3	-3	-2	-1
Printer	0	1	1	2	2	1	2	2	2	1	3	3	3		3	2	2	1	0	0	0	0	0	2	-1	0	0	0	-1	-3	-3	-3	-2	-2	-1	-2
Math., Phys. Sci. Teacher	0	0	0	2	2	1	2	2	2	2	3	3	3	3		3	3	2	0	0	0	0	1	1	-1	1	1	0	-1	-2	-3	-3	-3	-3	-2	-2
Policeman	-1	0	0	2	1	0	0	1	1	2	2	3	3	2	3		3	2	1	0	0	0	0	0	-2	1	1	0	0	-1	-1	-1	-3	-3	-3	-2
Forest Service Man	0	1	1	2	2	1	2	2	2	2	4	3	3	2	3	3		1	0	0	0	-1	0	0	-2	0	0	0	-1	-2	-2	-3	-3	-3	-2	-1
Y.M.C.A. Phys. Director	-1	-1	0	2	0	-1	-1	-1	0	0	0	0	0	1	2	2	1		2	3	3	2	3	1	-1	0	1	-1	0	0	-1	0	0	-1	-1	-2
Personnel Director	-2	-2	-2	0	-2	-2	-2	-1	-1	0	-2	0	-1	0	0	1	0	2		3	3	2	2	0	1	2	2	0	0	1	0	1	0	0	-1	0
Y.M.C.A. Secretary	-1	-2	-2	0	-2	-2	-2	-3	-2	-1	-1	-2	-1	0	0	0	0	3	3		4	3	3	1	0	1	2	0	0	1	0	2	0	0	-1	-2
Soc. Sci. H. S. Teacher	-1	-2	-2	0	-2	-2	-3	-3	-2	-2	-1	-2	-1	0	0	0	0	3	3	4		4	3	1	0	1	2	0	1	1	0	2	0	0	0	-2
City School Supt.	0	-1	-1	0	-1	0	-2	-2	-2	-2	-2	-2	-2	0	0	0	-1	2	2	3	4		3	1	1	0	0	-1	0	0	0	1	1	-2	0	-1
Minister	1	0	0	1	0	0	0	-1	0	-2	-1	-1	0	0	1	0	0	3	2	3	3	3		3	0	0	0	3	-1	-1	-1	0	1	1	1	-2
Musician	2	2	2	2	2	1	1	0	1	-2	0	0	0	2	1	0	0	1	0	1	1	1	3		0	-1	0	-3	-1	-2	-1	-1	1	0	1	-3
C. P. A.	0	0	0	-1	-1	0	0	0	0	0	-2	-2	-2	-1	-1	-2	-2	-1	1	0	0	1	0	0		1	0	0	1	0	0	0	1	2	1	0
Accountant	-3	-2	-3	-1	-2	-1	-1	0	0	2	0	0	0	0	1	1	0	0	2	1	1	0	0	-1	1		4	3	2	1	2	0	-1	-2	-3	0
Office Man	-3	-3	-3	-1	-2	-2	-2	-1	-2	1	0	0	0	0	1	1	0	1	2	2	2	0	0	0	0	4		3	3	1	1	1	-1	-2	-3	0
Purchasing Agent	-3	-2	-2	-2	-2	-2	-1	0	-1	2	0	0	0	0	0	0	0	-1	0	0	0	-1	3	-3	0	3	3		2	2	1	1	-1	-2	-3	2
Banker	-3	-3	-3	-2	-2	-2	-3	-2	-3	0	0	-1	-1	-1	-1	0	-1	0	0	0	1	0	-1	-1	1	2	3	2		2	2	2	0	0	-1	0
Sales Manager	-2	-3	-3	-2	-3	-3	-3	-2	-3	0	-2	-2	-2	-3	-2	-1	-2	0	1	1	1	0	-1	-2	0	1	1	2	2		4	4	2	1	0	2
Real Estate Salesman	-1	-2	-2	-2	-3	-3	-3	-3	-4	-1	-2	-2	-2	-3	-3	-1	-2	-1	0	0	0	0	-1	-1	0	2	1	1	2	4		4	1	0	0	2
Life Insurance Salesman	-1	-2	-2	-1	-2	-3	-4	-3	-4	-1	-3	-3	-3	-3	-3	-1	-3	0	1	2	2	1	0	-1	0	0	1	1	2	4	4		2	2	0	1
Advertising Man	2	0	0	-1	-1	-1	-1	-2	-2	-3	-3	-2	-3	-2	-3	-3	-3	0	0	0	0	1	1	1	1	-1	-1	-1	0	2	1	2		3	3	1
Lawyer	1	0	0	0	0	0	-1	-2	-1	-3	-3	-2	-3	-2	-3	-3	-3	-1	0	0	0	2	1	0	2	-2	-2	-2	0	1	0	2	3		3	0
Author-Journalist	4	2	2	0	1	1	1	0	0	-3	-1	-1	-2	-1	-2	-3	-2	-1	-1	-1	0	0	1	1	1	-3	-3	-3	-1	0	0	0	3	3		0
President - Mfg. Concern	0	0	0	-1	0	-1	0	0	0	1	1	0	-1	-2	-2	-2	-1	-2	0	-2	-2	-1	-2	-3	0	0	0	2	0	2	2	1	1	0	0	

Form 2. Computational form for "Highs", "Lows", and "Highs versus Lows".

CASE NO._____

	HIGHS:	LOWS:	HIGHS VS. LOWS:
PREDICTIVE INDEX:	_____	_____	_____

Form 3. Computational form for "consistency" scores.

CASE NO._____

AREA	ITEMS	L	I	D	MODE	MODAL SCORE	DEVIATION SCORE
Computational							
Sales (Selling)							
Social Service Work							
Biological Sciences							
Aesthetic							
Manual Worker							
Supervisor							
Clerical							
TOTALS							

Predictive Index: H_____ Z'_{1-2}_____

L_____

H vs L_____ Z'_{1-3}_____

Form 4. Computational form for "Similarity" scores.

CASE NO._____

FIELD	: LIKES	: DISLIKES	: LIKES FROM 34
Occupations			
School Subjects			
Amusements			
Activities			
Possession of Desirable Traits			
TOTALS			

SIMILARITY SCORE Dislikes plus Likes from 34	

Form 5. Summary of computations.

NAME_____ CASE NO._____

Z'_{1-2}_____ Z'_{1-3}_____

Hi_____ Lo_____ Hi vs Lo_____

Modal Score_____ Deviation Score_____

Similarity Score (25 year olds)_____

PREDICTION OF THE PERMANENCE OF INTERESTS

by

JAMES LUKE SMITH JR.

B. S., Kansas State College
of Agriculture and Applied Science, 1955

AN ABSTRACT OF A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1956

This study investigated three major hypotheses concerning the prediction of the stability of vocational interests as measured by the Strong Vocational Interest Test. These hypotheses were:

1. A stable Strong Vocational Interest Profile (one that is reliable over time) will show a higher order of consistency among its scales than an unstable one. Three specific sub-hypotheses were derived from this.

a. The relationships among the five highest occupational scores will be of a higher order for stable profiles than for unstable ones.

b. The relationships among the five lowest occupational scores will be of a higher order for stable profiles than for unstable ones.

c. The relationships among the five highest scores, on one hand, and the five lowest, on the other, will be of a greater negative magnitude for stable profiles than for unstable ones.

2. Consistency of response to a core of similar items will be related to profile stability.

3. Similarity in item response to 25-year olds as opposed to 15-year olds will be more characteristic of stable profiles than of unstable ones.

The sample used for this study was 176 of the 181 students used by Stordahl. These students were originally tested on the Strong in 1949 as high school seniors. Stordahl retested them in 1951 as college sophomores. For a portion of the experiment, 116 of these students who were tested for a third time in 1953 by Hoyt composed the sample.

Two criteria were employed. The first was the z' equivalent of the rank order correlation between 1949 and 1951 profiles. The second was the same measure applied to 1949 and 1953 profiles. Earlier work had established the z' index as a valid indicator of interest stability.

Six predictive indices were devised on the basis of the hypotheses. Statistical tests of these hypotheses were made by two methods:

1. Correlational analysis.

2. Analysis of mean score differences.

In the correlational analysis, each of the six predictive indices was correlated with both criteria. For the second method of testing the hypotheses (analysis of mean score differences), groups of "High", "Average", and "Low" stability were defined on the basis of z' scores. Means and standard deviations for each of the three groups on each of the six indices were computed, and t tests of the significance of the difference between the means of the "High" and "Low" groups were run.

Within the limits of the sample used, the following conclusions seem warranted:

1. The first hypothesis was supported. A statistically significant relationship was established between each of the three predictive indices and both criteria. This conclusion held true regardless of the method used to test the hypothesis.

2. Despite the general confirmation of the first hypothesis, the magnitude of the relationship between the predictive indices and the criteria was unimpressive. Coefficients of correlation varied from .20 to .37, Overlap varied from 18 per cent to 43 per cent.

3. Of the indices, the one designed to test sub-hypothesis (c) was consistently more closely related to the criteria than either of the other two.

4. Optimally combining the three predictive variables through the multiple correlation method did not substantially increase their relationship to the criteria.

5. Stability over a 4-year period was predicted slightly more accurately than stability over a 2-year period.

6. The second hypothesis was largely unsubstantiated by the data. In general, the trend of the results was in line with the hypothesis, but statistical significance (and practical significance) was lacking.

7. No support was found for the third hypothesis. Similarity in response to 25-year olds as opposed to 15-year olds apparently had no more significance for interest stability in this case than it did in the case of Strong's Interest Maturity scale.

Several implications for further research were discussed.


STATE COLLEGE OF AGRICULTURE AND APPLIED SCIENCE
LIBRARY
MANHATTAN
KANSAS