



eCOMMONS

Loyola University Chicago
Loyola eCommons

Dissertations

Theses and Dissertations

1963

The Reliability of the Loyola Language Study: Its Relationship to Values, Interests, and to Group Mindedness

John V.P. Stewart
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

 Part of the [Psychology Commons](#)

Recommended Citation

Stewart, John V.P., "The Reliability of the Loyola Language Study: Its Relationship to Values, Interests, and to Group Mindedness" (1963). *Dissertations*. 712.
https://ecommons.luc.edu/luc_diss/712

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
Copyright © 1963 John V.P. Stewart

THE RELIABILITY OF THE LOYOLA LANGUAGE STUDY:
ITS RELATIONSHIP TO VALUES, INTERESTS,
AND TO GROUP MINDEDNESS

by

John V. P. Stewart

A Dissertation Submitted to the Faculty of the Graduate School of
Loyola University in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

January

1963

LIFE

John V. P. Stewart was born in Sterling, Kansas, May 17, 1907.

He was graduated from Hutchinson High School, Hutchinson, Kansas, in June, 1925 and from the University of the South, Sewanee, Tennessee, June, 1930, with the degree of Bachelor of Arts.

He then entered the Theological School of the University and graduated in 1934 with the degree of Bachelor of Divinity.

From 1934 to 1952, the author served as a clergyman of the Protestant Episcopal Church. Five years of this time was spent as a Chaplain in the United States Navy and in the Marine Corps.

In 1952, having become a convert to the Catholic Church, he resigned from the Protestant Episcopal ministry and entered the Graduate School of Loyola University of Chicago to begin studies in clinical psychology. The author received

the Master of Arts degree in 1956. While continuing his studies, he was employed as a Psychologist in the Loyola Center for Psychological Services until 1959 when he entered private practice. During his full time private practice he has worked very closely with Dr. James H. Hutton, an endocrinologist, and was supervised by the late William J. Devlin, S.J., M.D., Ph.D.

TABLE OF CONTENTS

Chapter	Page
I. Introduction.	1
II. Review of the Literature on Association Tests	11
III. Procedure	24
IV. Discussion of Results	33
V. Summary and Conclusions	48
Bibliography.	52

LIST OF TABLES

Table	Page
I. Rank Difference Correlations between Allport-Vernon-Lindzey Scale of Values and the Loyola Language Study for 17 Men, Tested in 1954, Retested in 1958	34
II. Rank Difference Correlations between Allport-Vernon-Lindzey Scale of Values and the Loyola Language Study for 20 Women Tested in 1954, Retested in 1958	35
III. Rank Difference Correlations between Kuder Preference Record Interest Scores and the Loyola Language Study Scores for 17 Men Tested in 1954, Retested in 1958	36
IV. Rank Difference Correlations between Kuder Preference Record Interest Scores and the LIS Scores for 20 Women Tested in 1954 and Retested in 1958	37
V. Mean Scores for Men, and Differences between Group I and Group II on the Kuder Preference Record. Group I are the High-Activities and Group II are the Low-Activities Persons. .	43
VI. Scores on the Allport-Vernon-Lindzey Scale of Values made by Men. High Activities Persons are Group I and Low Activities are Group II.	44
VII. Scores on the Kuder Preference Record made by Women. High Activities Persons are Group I, and Low Activities are Group II	45
VIII. Scores on the Allport-Vernon-Lindzey Scale of Values made by Women. High Activities are Group I, Low Activities Group II	45
IX. Fisher's Exact Probabilities Test, to discover the Significance of the Differences between the Active and Non-Active Scores of Men on the Kuder Preference Record.	46a
X. Fisher's Exact Probabilities Test to Discover Significance of Difference between Active and Non-Active Men on Two Scale of Value Scores (Theoretical and Aesthetic) on the Allport-Vernon-Lindzey Scale of Values	47

Table

Page

XI. Fisher's Exact Probabilities Test, to Discover the Significance of the Differences between Active and Non- Active Women's Scores on the Kuder Preference Record (Scientific and Persuasive).	47
---	----

CHAPTER I

INTRODUCTION

The Loyola Language Study is a controlled word association test. It makes use of eighty of the one hundred stimulus words used by Kent and Rosanoff in their famous test of free association. It is called a controlled association test because the instructions given to the subject limit the kind of response to be given, not in the manner of the traditional control but in an entirely different manner. Whereas the older controlled association tests would require the subject, after sufficient reflection, to give a subordinate or supraordinate to the concept contained in the stimulus word, the Loyola Language Study asks the subject to pause to think what he believes the majority of persons would judge to be the right response to make to the particular stimulus.

The rationale of the test involves the concept of communality of thought as measured by a person's ability to give associations shared by the majority of the population. It also measures the individual's awareness of that communal element and his own conformity to, or deviation from, "normal thought." This approach then is entirely different from the traditional free association response to stimulus words. In her article, "The use of Free Association in the Objective Measurement of Personality," Goodenough (1942) wrote:

It is a well recognized principle of the free association technique that the subject must be told to give the first associated word that comes to his mind. 'Considered' responses are of less significance

because, in the process of taking thought, ideas of social acceptability, of intellectual values rather than basic attitudes, emotions and interest patterns, are likely to determine the response.

This very "process of taking thought," which Goodenough considered less significant, is the significant aspect of the Loyola Language Study. Because of this process, past experiences, reasoned evaluations, deliberations, choice and other factors of ego control may be called upon in responding to the word stimuli. It is this which gives it originality as a controlled word association test. The Loyola Language Study, henceforth to be called the LLS, is then neither a free association test nor is it a logically controlled association test. Rather, it is a test of the ability of the subject to determine through some means or other, what he thinks is likely to be the response that would be most frequently given by the majority of the population.

The LLS is the outgrowth of research undertaken in 1953 by the late Rev. Louis B. Snider, S.J. and Olof Johnson at Boston State Hospital, Boston, Massachusetts. The method of control is unique and was hit upon after a series of preliminary investigations in which Snider and Johnson had found that the subjects did not always give back the first word that came to their minds when they were asked to find the 'response that belongs' to a particular stimulus word. The first reports on their findings were to the staff of psychiatrists, psychologists and social workers at the Boston State Hospital. From the first, the little group of persons that was allowed to share the secret of Snider and Johnson, seemed convinced both theoretically and practically, that the patients in the hospital would not do as well as normal subjects, when asked to respond according to this new association technique. Incidentally, the earliest form of the test went by the name of "What would most People think of Test." From

the beginning the researchers agreed that the new form of instructions should be standardized. It was noticed that many people would spend too much time trying to decide "what most people would think" and complained that they could not possibly find the answer. Consequently, the form of instructions actually used on the printed booklets was the outcome of months and months of preliminary investigation and experimentation. As can be seen in the sample at the end of this dissertation, subjects taking the test are told to take as much time as they need to think about the word which seems to "go along" with each printed stimulus word. They are told not to write down just any word that comes to mind but only the ONE WORD that they think most other people would be likely to think of.

The LLS became the property of Loyola University of Chicago, after Olof Johnson found it necessary to abandon work on the project, and turned all of his data over to Father Snider. Just before his untimely death, Snider travelled back to Boston to collect some needed data; he found Mr. Johnson much interested in the results, but not then taking an active part in the project.

Since 1954, extended research has been conducted on the LLS. In 1954 the LLS was administered to a group of college freshmen both men and women, entering Loyola University of Chicago; and a research project at the M.A. level was carried out by the present investigator. This project will be described in Chapter II. It seemed reasonable to assume that by administering the same test to a group of graduating seniors who had already taken the test in 1954 as freshmen, it would be possible to secure adequate data for the determination of the long-term reliability of the test. This is one of the most important of

several purposes of the present research. V. V. Herr (1958) has already reported an estimation of the short-term reliability and his findings compare favorably with any of those found in the literature for free association tests, though specific reports on reliability are rather scarce. (*)

As a result of the examination of the literature it seemed not only desirable, but even necessary, to make a study of the reliability of the LLS over a much longer period of time, even as much as four years, since such an investigation seems not to have been carried out previously involving the use of other association tests.

In the Snider-Johnson study as in the present one, men and women are studied separately. The reason for this is the fact which other investigators pointed out from the very beginning, namely that top frequency responses appeared to be somewhat different for men as compared with women.

An unpublished study by Dinello (1958), indicated that there were trends toward significant differences in scores on the LLS made by various professional groups. By trends are meant low but persistent correlations in the same positive direction. In most of the word association reports found in the literature, there was evidence of serious interest in those trends or small tendencies. In fact, many of the tentative conclusions from the very beginnings of word association studies were based on such persistently recurring relationships.

*The earlier data for the LLS reliability studies were collected by Drs. Eugene Albrecht and Robert Nicolay, using their classes of over 200 students each as experimental subjects. The data are on file in the Lake Shore Campus Psychology Laboratory.

Florence Goodenough (1942) raised many problems regarding word association and yet she still felt that some form of the free association method offered one of the most promising approaches to the study of personality differences because of the noted tendencies to make various responses. Also in the Kent-Rosanoff (1910) study, the particular interest of the experiment lay in its affording a precise measurement of the tendency of a person to respond along the same lines as his fellows, or in ways peculiar to himself.

The interest in these trends or tendencies has been kept alive because of the important tendencies that have already been discovered, involving the use of the ILS. Trends towards significant differences in scores on the ILS made by various professional groups seem to provide clues too important to ignore. We repeat: By trends we mean only persistent low correlations in the same direction which usually turn out to be positive and significant when more or larger groups are used. In the study of personality variables one frequently sees a good deal of meaning in these recurring relationships.

In Dinello's study there is the suggestion that communality of thought was more characteristic of some groups than of others. The groups of men who scored better on the ILS tended to be those whose occupations required much contact with people; for example, salesmen and managerial persons. Since in this study only men were tested, it appears that further research in this area, including women, might be profitable. Consideration of these findings and other similar ones led to the hypotheses studied in the present investigation. It is unfortunate that Dinello's sample was so small. One understands the reason for this when he reflects that Dinello was obliged to match the several groups used for comparison on the two variables of age and education, since

these had been shown to contribute very much to the variance of the ILS scores.

The reasoning from which hypotheses were developed is somewhat as follows: there are valid and reliable tests of interest such as the Kuder Preference Record, the Strong Inventory, and others. If, as certain studies have appeared to indicate, some persons are more interested in occupations in which they would deal with people than they are in studying abstract science which might lead them to the quiet seclusion of a research laboratory; then this interest in occupations requiring contact with people, with all of the associated obligations and responsibilities, might be expected to show up in a higher communality score on the ILS. This would be true since occupations requiring successful interpersonal relationships require varying degrees of the capacity for empathy. We here look on empathy as the capacity to identify with another and experience vicariously what he experiences. For reasons to be explained later, the Kuder Preference Record was the instrument chosen to test the hypothesis. Interest levels measured by the Kuder are: Outdoor, Mechanical, Computational, Scientific Persuasive, Artistic, Literary, Musical, Social Service, and Clerical. From the trends noted in Dinello's study, it might be hypothesized that persons with high Persuasive and Social Service scores would also make better communality scores on the ILS, while those with high Scientific, Clerical, and Computational interests might make poorer communality scores on the ILS.

Similarly there is an instrument which measures the systems of Values of men and women, namely the Allport-Vernon-Lindzey Scale of Values. Validity and reliability figures for this Scale and for the Kuder Preference Record are given in Chapter II. The Scale of Values aims to measure the relative

prominence of six basic interests or motives in personality, these being the Theoretical, Economic, Aesthetic, Social, Political and Religious. This classification is based directly upon the work of Edward Spranger (1924) a work which upholds the view that men's personalities are best known through their systems of value or evaluative attitudes. A broad scope would seem to be provided by such a scale for investigation of the relationship of communality of thought to personality insofar as evaluative attitudes are indicative of personality traits.

Jenkins (1959) studied the effect of the instruction to give "popular" responses upon the word-association behavior. He concluded that the "popular set" markedly increases the number of top frequency responses, and that, judging from intercorrelations between the test with standard instructions and the test with the "popular set," a person's score on one test does not contribute to his score on the other. Finally he adds: "the instructors involved suggest from their knowledge of the Ss that high "gain scores" are related to what might be called "social sensitivity." In a study concerning Predictive Empathy and the Study of Values (Halpern, H.M. 1957) found a positive correlation between empathy and social values and a negative correlation between empathy and Aesthetic values. The sample was composed of women. Further study using both men and women might be profitable. Halpern's study lends support to the decision to study the following hypothesis: there might be a relationship, for example, between the possession of strong social values and the empathetic capacity whose strength is measured by the LLS. Similarly it would seem that a person whose high value score was in politics should score better on the LLS than one whose strongest value score was in aesthetics. The

reason for the difference would be that one whose value score was high in politics would be likely to be more aware of how other people think, than would a person with high value on aesthetics.

These two hypotheses, first that there should be a relationship between certain occupational interest level scores and scores on the ILS, and second, that we might expect to find a relationship between certain values and ILS scores, led to the decision to administer the Kuder Preference Record and the Allport-Vernon-Lindzey Scale of Values to the college seniors at the time they came to take the re-test ILS.

Furthermore it was argued that persons who join many fraternities or who generally take part in more extra-curricular activities than fellow class members, might be expected to have better communality scores on the ILS. As a result of this consideration it was decided to dichotomize the total group of those seniors who were to take the repeat ILS into the active and the non-active groups. It would then be possible to compare the tendencies of these two groups to score differently on the ILS, on the different scales of the Kuder Preference Record, and on the separate scales of the Allport-Vernon-Lindzey Scale of Values.

Hence the following were decided upon as the objectives and goals of this research. There was sought an index of reliability of the ILS over a period of four years. This will be called its long-term test-retest reliability. Second, the possible relationship between communality scores on the ILS and the high or low interests in various occupations, as indicated on the Kuder Preference Record, will be determined. Next, the possible relationship will be sought between communality scores on the ILS, and certain values, as measured by the

Allport-Vernon-Lindzey Scale of Values. Fourth, possible relationships will be sought between ILS scores and what might be called a tendency to join in social activities. This tendency will be evaluated by ascertaining the frequency with which students engage in extracurricular activities while in college. The active ones would presumably be more likely to have high communality scores and the non-active the reverse. Similar relationships will be sought between the frequency of joining and scores on the Scale of Values, and between frequency of joining and scores on the Kuder Preference Record.

The importance of discovering these relationships arises because of the fact that the ILS is already known to discriminate between normals and mental patients belonging to the category of schizophrenia. These actually are known to have very poor communality scores. It is also known to be somehow related to certain personality variables that involve sociability. These relationships are not strong, but they are persistent and reappear whenever one takes the time to make the investigation. It is true that in each of the previous studies, the samples were small. Yet the present investigation aims at adding a few more samples to those already in existence. After all it is no easy matter to collect huge samples of persons or large numbers of patients who are willing to take these tests.

Since there had been many samples used in testing the stability of the ILS scores over a period of time lasting as much as four months, it was now decided to check the stability, by a test-retest correlation, after a span of time extending from the first to the fourth year of college. Should the correlations between the first score and that obtained four years later, agree with those obtained when the time interval was only four months, there would

be additional evidence for the stability of the test scores. Moreover, should the IIS scores obtained in 1954 as well as in 1958 show consistent correlations with the personality variables mentioned above, there would also seem to be indicated a rather uniform tendency for the communality of thought scores to reveal personality traits.

This study is one of several others designed to find out precisely what it is that the IIS measures. Upon its completion the word association study begun by Father Snider, whose untimely death at the age of 43, was very unfortunate, and carried on by the Reverend Vincent V. Herr and others at Loyola University, promises to be one of considerable value in the history of word association studies.

CHAPTER II

REVIEW OF THE LITERATURE ON ASSOCIATION TESTS

Highly optimistic predictions regarding word association tests as of possible value in the study of personality have been plentiful since the days when Dr. Carl Jung (1910) was writing articles and giving lectures on the "Association Method." In view of such optimism, and the favorable attitudes toward association tests expressed by many who are generally admitted to have played major roles in the development of Psychology as a science, as, for example, Galton, Wundt, Cattell, Kraepelin and F. Goodenough, it is surprising to find how little research evidence has accumulated regarding the reliability and validity of such tests. Herr (1958) also noted in his article on the ILS, the scarcity of reliability reports on word association tests in the literature.

While there are reports on controlled tests in the literature, they are far outnumbered by the free association tests developed. There is no association test reported in the literature using a control such as that employed in the ILS and described on page 3.

When, in 1883, Galton first began working with word association, Wundt very soon saw the possibilities and began working with the method experimentally. Interest in word association techniques was renewed by Jung in 1910. He deserves much credit for a renewed arousal of interest in the use of word association as a guide in personality study. Kent, G. H., and Rosanoff, A. J., (1910) published the first notable piece of research on the relation of free

association to insanity. Rosanoff, Isabelle and Rosanoff, A.J. (1913) published research dealing with children. Eventually, Rosanoff (1927) brought the then existing research together into one volume. O'Connor (1928) did an extensive item analysis of the Kent-Rosanoff test data which he had collected from a sample of male factory workers, and he compared briefly the popular responses with those of the Kent-Rosanoff mixed adult sample of 1910. He went no further than to report:

"The common response to a stimulus word is, therefore, a scientific reality in the sense that it is rediscoverable by different workers in a new laboratory after a lapse of 15 years and the identity of a common response is more reproducible than its frequency of appearance."

Malamud (1946) working with Maller's Controlled Association test reported both validity and reliability satisfactory, even for individual diagnosis using the test for a screening device. But his actual figures are not reported. Conrad, H. A. and Harris, D. (1931) in their study of Free Association Method and the Measurement of Adult Intelligence, conclude that the reliability of scores derived from their twenty hard words, though not computed, might be inferred to be fairly satisfactory. Here again, actual figures are not reported. Tresselt and Leeds (1955) compared word association responses collected in 1952 on a small sample, heterogeneous with respect to residence and occupation, with what they referred to as the original norm established on the Kent-Rosanoff in 1927. (The 1927 norms are the same as the 1910 norms; so actually, the 1910 norms are the "original norms.") They reported a decided increase in primary responses which were responses of opposites such as Dark-Light, Black-White, co-existing pairs, such as King-Queen, Salt-Pepper. Doerkin, (1956) reported a study of ten high frequency stimulus response pairs from the Kent-Rosanoff

list in various normative collections; he found that the frequency of these responses appeared to be increasing steadily since 1910. It remained for Jenkins and Russell (1960) to

"make clear both the magnitude and the nature of the changes revealed by the comparisons of word association norms from 1910 to 1952 and to argue that these changes reveal a systematic and important trend in responding to free-association tests."

In summarizing, these latter investigators reported the development of three hypotheses from their comparison of the 1927-1952 responses. These hypotheses were studied with respect to some of the more prominent collections of free association data, the collections being: 1. the 1910 and 1927 Kent-Rosanoff norms; 2. the 1928 O'Connor norms; 3. the 1951 Keene norms; and 4. the Jenkins and Russell Minnesota norms. The first hypothesis was verified, that is, it appeared that there is a general tendency for the frequency of popular responses to increase with the passing of time. This was, however, not a certain conclusion, owing to the fact that other factors, such as the method of testing, may have somewhat contributed to this change with time. Keene and others have mentioned as being significant factors to be controlled, the manner of giving the tests, the personality of the tester and the like. These would all have to be taken into account in judging the results. These variables would not be likely to enter into the administration of the ILS, since the instructions are printed on the test booklets, and are only emphasized by the person administering the test, to make sure they are fully understood.

The second hypothesis of Jenkins and Russell, namely that there were some changes in the responses with time, was also confirmed; and that the responses that had top frequency tended to be more stable than the others. They further noted that abstract responses, and in particular, superordinates, tended to

decrease in popularity during the time covered by their experiments. This was their third hypothesis. Their other hypotheses had to do with the general change in the test-taking attitudes; perhaps these changes would account for the increase in popular and decrease in abstract responses; perhaps also the changes in meaning of the stimuli over a period of time might have accounted for the changes, but the investigators could not well understand why the changes should have gone in that particular direction, merely from changes in the meaning of the stimulus words.

Their study showed that free association responses are therefore subject to the influences of cultural and local factors. And thus the norms set up by Kent and Rossnoff in 1910 and in 1927 would be said to have some validity, at least for contemporary Americans, but this would not be absolute.

While the 80 words of the LLS were selected from the Kent-Rosanoff list, the total test is not comparable to that used by the original authors. The reason is that in the LLS there are completely different instructions. Moreover new and distinct norms were established for the LLS by the workers at Loyola, which have proved to be very stable over the years since 1952. Undoubtedly there will be some changes with time; Loyola workers have already catalogued the changes that are to be expected in three major geographical areas of the United States.

It is likely also that the meanings of some of the words will change with time, and some words may also become obsolete. These changes will have to be expected, and taken into account by the later research workers. Till now, there has been a relatively high stability of responses for a given sex and age-education group. It should be born in mind that in the LLS the subjects are

not asked to act spontaneously, but to reflect and try to ascertain by thorough evaluation processes which might be the response that would be selected by most people. It has been found by many investigators that with this instruction the subjects do not give back the same responses which they would give if told to give the first response that comes to mind "as quickly as possible." Again the writer of this dissertation knows from personal experience that when the experimenter omits the words "as quickly as possible" he may find different results. For this reason the instructions for association tests should always be carefully standardized.

In Chapter III, more details will be given in order to show the importance of this matter of instructions.

M. L. Sorenson and H. D. Carter (1940) found that association lists of at least 48 words were necessary to secure reliability coefficients of .80 or better in measuring twin resemblances in community of response. In an unpublished doctoral dissertation, Keene, (1951) reports the finding of a general and constant tendency for an individual to give the same type of response on repeat tests. This was considered justification for further study in this area. The test was repeated using the original group of 500 for two successive years. Most of the correlations between the scores for the different years were around .65 or better. Some for the females dropped slightly below .50. Keene further reports:

"Since the re-tests were given by the same individual the impact of the tester and his specific method of test administration appeared to have more influence on test scores than the type of frequency tables used."

On the other hand, an unpublished study by Clousing (1927) indicates the negligible effect of method (i.e. oral or written) as long as both methods ask

that responses be given quickly; no differences were found between the group written form of the test and the individual oral administration, in the frequency of primary, secondary, or tertiary responses. Keene (1951) also asserts that there is little difference between the two modes of testing, although he does not present his group data for comparison with his tables of individual data.

Literature reviewed on the preceding pages shows the lack of specific reliability reports on word association tests, and so emphasizes the importance of such an investigation on the LLS.

As mentioned on page 4, Herr (1958) did establish LLS retest reliability for time intervals as long as four months. In his study the test had been administered to the same persons twice, once at the beginning of a semester or of a quarter and once at the end. In his study a reliability coefficient of .49 was found, taking men and women together. However, when only the twenty-five items having top screening efficiency were used, the correlation rose to .55; studying men and women separately, the women scored slightly higher reliability than the men.

Del Vecchio (1957) found that the LLS does distinguish between schizophrenics and normals, and Herr (1958) showed the same would happen using any of three systems of scoring, i.e. standard scores, doubled-root frequency scores and median scores. Herr also reported that the screening efficiency of the test is increased by using only high validity items scored by standard scores. Such scores also permit of parametric statistical operations. Older persons tend to make poorer scores, though this effect tends to be counteracted by education. Intelligence, apart from age and education, is not a significant factor. This

confirms the findings reported by Stewart (1956).

REVIEW OF LITERATURE ON VALUES AND INTERESTS

The Allport-Vernon-Lindzey Study of Values, "A Scale for Measuring the Dominant Interests in Personality," and the Kuder Preference Record were the instruments chosen to test two hypotheses mentioned in the first chapter on pages 6 and 7. The revised form of the Study of Values was used. Originally published in 1931, by Allport and Vernon, it was the first study of its kind. Certain improvements are offered in the revision without changing the basic purpose of the test or limiting its usefulness. Of importance in the present research is the fact that the revision of the Study of Values required the preparation of new norms. Since the ILS norms are contemporary, the revision with the new norms is especially attractive for the purposes of this particular study. The items of the Study of Values are designed to measure the following interests or values: Religious, Political, Social, Aesthetic, Economic and Theoretical. In the 1931 test, only in the Social values scale was the reliability questionable when diagnostic efficiency was tested by the internal consistency method. In the revision, the definition of the "Social" value has been made more specific. Correlations between the old form and the revised form (taken two weeks apart) are all significantly high. These correlations were based on a sample of 50 male college students; the results might have been more accurate had a sample of 50 women been included. Work with the ILS has revealed a marked sex difference in response to stimulus words, and we might expect a sex difference in the relationship of the ILS to Values.

The reliabilities of the Study of Values obtained by various methods seem satisfactory, especially in view of the fact that each value is measured by only

twenty items. The internal consistency of the scale was determined by two methods, i.e., split-half reliability and item analysis. The mean reliability coefficient using a z transformation was .82 as opposed to .70 for the old form. The final item analysis carried out on a group of 780 subjects of both sexes from six different colleges, shows a positive correlation for each item with the total score for its value, significant at the .01 level of confidence.

The table presented in the Manual gives r 's obtained for a heterogeneous sample of 100 males and 100 females. There is a positive relationship between social-religious values, and between economic-political; less association is indicated between theoretical-aesthetic.

It is stated in the Manual that the old form of the test has been successfully used in a wide variety of psychological researches, and that the Revised Form should have similar uses. "Wide variety" here means at least forty-five different research projects. In view of the significantly high correlations between the old and the revised forms, and the improvements made by the authors, the Revised Form was used in the present study.

A brief explanation of Spranger's Types of Men is given in the Manual and presented here, since it is expected that the subjects with high ILS scores will have higher Values scores in those areas where an empathetic capacity would seem to be required. It should be helpful then to understand upon what types the various Values are based.

- (1) The Theoretical. His dominant interest is discovery of truth, his interests are empirical, critical and rational. Necessarily an intellectualist, his chief aim in life is to order and to systematize his knowledge.
- (2) The Economic. This man is characteristically interested in what is useful. His interests develop to embrace the practical affairs of the business world, and the accumulation of tangible wealth. This type is thoroughly 'practical' and conforms well to the prevailing stereotype of the average American business man. . .this type is more likely to be interested in

surpassing people in wealth than in dominating them, or in serving them (social attitude.)

- (3) The Aesthetic. He places highest value upon form and harmony. Every experience is judged from the standpoint of grace, symmetry, or fitness. This man considers truth equivalent to beauty. He may be interested in persons, but not in the welfare of persons; he tends toward individualism and self-sufficiency.
- (4) The Social. For this type, the highest value is love of people. It is the altruistic or philanthropic aspect of love that is measured. (Here the authors have restricted Spranger's type.)
- (5) The Political. Interested primarily in power. He may not be a politician, but whatever his field, his manner and methods will betray him as a *Machtmensch*.
- (6) The Religious. According to Spranger, this man's highest value might be called unity. He is one whose mental structure is permanently directed to the creation of the highest and absolutely satisfying value experience. He may be a mystic, as an ascetic, withdrawing from life to find unity through self denial and meditation; or he may be an 'immanent mystic' finding his religious experience in active participation in life. Spranger does not mean that any given man belongs exclusively to one or another of these types of values; nor is the Study of Values based upon such an exclusive idea.

Both the Strong Vocational Interest Blank and the Kuder Preference Record were considered as instruments for testing the hypothesis that persons interested in occupations bringing them into close contact with people might be expected to make higher scores on the ILS than those who were more interested in occupations in which inter-personal relationships would be at a minimum. While both tests are widely used, the Kuder was chosen on the basis that it is easiest to comprehend, easiest to mark, and mechanically attractive.

The basic premise in the development of the Kuder Preference Record is that people do best when they are interested in what they are doing and enjoying themselves while doing it. It is taken for granted that a high interest score in a particular area, or in more than one area, does not necessarily mean that the person making such a score will be successful in that area. For example, he may have a high scientific interest, but a rather low intelligence level. He might wish to be a doctor; his high scientific interest would be compatible with such

a wish but his intelligence level would make it inadvisable to encourage him. The Kuder Preference Record is intended primarily for vocational guidance. The author of the test began by listing a large number of activities from work, school, and recreation. An analysis of correlations was then made to determine which of these interests were found together. Seven areas of interest are included in the test, then two additional scales were added which partly overlapped two of the original ones. The final profile describes the subject in terms of his interests in outdoor, mechanical, computational, scientific, persuasive, artistic, literary, musical, social service, and clerical activities. Each Kuder item has the form of a forced choice among three possible activities, the person being required to select the one he likes least and the one he likes most. The resulting profiles are interpreted on their face value. A person who shows high interest in Clerical and Computational would be expected to enjoy positions demanding such activities.

The following is condensed from the Kuder administrator's manual, 1960 revision. The ten broad areas in which preferences are measured by the Kuder Preference Record are:

0-Outdoor: Indicates a preference for work that keeps one outside most of the time, usually dealing with animals and growing things.

1-Mechanical: Indicates a preference for work with machines and tools.

2-Computational: Indicates a preference for working with numbers.

3-Scientific: Indicates preference for discovering new facts and solving problems.

4-Persuasive: Indicates a preference for meeting and dealing with people, and promoting projects or things to sell.

5=Artistic: Indicates a preference for doing creative work with one's hands. It is usually work that has 'eye appeal' involving attractive design, color and materials.

6=Literary: Indicates a preference for reading and writing.

7=Musical: Indicates a preference for going to concerts, playing instruments, singing or reading about music and musicians.

8=Social Service: Indicates a preference for helping people.

9=Clerical: Indicates a preference for office work that requires precision and accuracy.

A verification or V scale is included as an accuracy check, intended to identify persons who may have responded carelessly. This scale was developed by assembling items which yielded extreme response splits. V-scale items were chosen on the basis of item counts for several different groups, which included high school and college students as well as adults. The total number of cases involved was between 2,500 and 3,000, and items were included in the scale if a single response or combination of two responses (to a group of three activities) was marked by at least 90% of the subjects in each group studied. In these studies, the number of persons who obtained scores of less than 38, which was the cutting point for validity, varied from one to five percent from group to group. Computation indicated that chance scores on the V scale would have a mean of 25 and standard deviation of 3.65. The author's empirical check of these figures on a group of 109 subjects who were asked to mark the test carelessly yielded a mean V-score of 24.81, and a standard deviation of 3.70. If a subject's score is not within the range of 38 - 44 inclusive, there is some reason to doubt the value of his answers. All of the subjects in the present

study met this requirement.

According to Cronbach, Lee J. (1949) the logical validity of the Kuder is high. Triggs, F. O. (1947) reports interests measured by both the Strong and Kuder tests have correlations near .60. The correlations in themselves are some evidence that the tests are valid, since tests made by very different techniques give fairly similar pictures of the individual.

When the reliability of interest scores is studied by immediate retests, coefficients are high, as would be expected from the large number of items used. But a more important question is the stability of the scores. While there is not as much data available for the Kuder as for the Strong, one study, Traxler, A. E. and McCall, W. C. (1941), of adults tested 15 months apart showed correlations as high as .93 (Musical) and as low as .61 (Social Services). The median correlation was .83. Crosby, R. C. and Wisnor, A. L. (1941) found correlations between scores on the Kuder and self-estimated interests to be as follows: Scientific, .48; Musical, .58; Social Service, .39; Persuasive, .62; average, .52. This is significant as evidence that self-reports in response to varied items as in the Kuder, give a picture that differs from responses to occupational labels. Cronbach (1949) states that:

"The evidence that interest scores do discriminate between men in various occupations is partial evidence that interests are a sound basis for guidance. Both the Strong and Kuder tests have been studied sufficiently to verify that the majority of successful men in an occupation have corresponding scores on the interest tests."

An unpublished study at the M.A. level by Stewart, J. V. P. (1956) included a careful review of word association literature from the pioneer work of Galton in 1883 to 1955. No form of the word association technique employing the method of control developed by Father Snider and Olof Johnson, described in Chapter I,

was found. Furthermore the majority of the reports in the literature were concerned primarily with the diagnostic features of word association tests. Also, as has been shown, very few researchers have been concerned with establishing reliability.

This research is investigating areas not before studied in word association research. Since the LLS does measure a person's ability to think as others think and probably indicates a capacity or lack of capacity for empathy, because of reasons set forth before, it should be of value as an instrument supporting the findings of such tests as the Strong Vocational Interest Blank, the Kuder Preference Record, the Allport-Vernon-Lindzey Study of Values and others. It may also indicate those who may be more inclined to join in social activity than others.

CHAPTER III

PROCEDURE

The procedure of Snider and Johnson in establishing their Boston norms has been described in Chapter One.

Father Snider had recognized from the beginning that norms would vary geographically, and it was at his suggestion that Stanek, Richard (1953) undertook the establishment of Chicago LLS norms, using standard scaled scores. Snider's standard scores were used throughout the entire present research. Following is a brief review of Stanek's procedure in establishing adequate norms for the stratified sample derived from the Chicago Metropolitan area.

Stanek's material was located and collected according to a modified method of stratified-random sampling. The population of the Chicago metropolitan area was analyzed to determine the proportions of the total population within each category of age, sex, and education. It was also necessary to adjust the census figures used for the analysis, since they included all children over five years of age, some illiterate adults, some mentally retarded etc. This adjustment applied particularly to the lower educational groups. Ages were divided into three twelve year units, i.e. 19 through 30, 31 through 42, and 43 through 54. The sample was distributed according to educational level for each age group, using four step intervals of three grades each, sixth grade through eighth grade, ninth through eleventh grade, twelfth through fourteenth grade, and fifteenth grade and higher. From the proportions of the population falling into these

categories it was possible to collect a sample corresponding approximately to the expected proportions of a population which is well above the level of the mentally deficient. Within each age group and educational level, an attempt was made to select individuals distributed through the actual age ranges and educational levels of each cell.

The first phase of the procedure was to determine the various proportions of the Chicago metropolitan population falling into the three categories of age, sex and educational level. The number of subjects in the Chicago sample was 800; appropriate proportions, corresponding to proportions of the population were calculated in order to obtain an approximately representative number of subjects in each category. The proportions were determined from the latest national census, 1950. The test being self-administering, was distributed in two ways: first by giving the test forms directly to various groups of adults to be completed at the time. These were such groups as factory employees during a rest period or some group in a training session. Second, by requesting individuals or small groups in various institutions and agencies throughout the area to give the test forms to their relatives and/or friends, to be mailed when completed. (As has been mentioned, the explanations are direct and simple, making the LLS an excellent instrument to be used when the situation is such that the subject cannot be present.) Many more tests were obtained at night school classes as well as from various clubs meeting at night.

In establishing norms for the general population, Chicago and Boston, males and females, they were treated as four separate categories. A frequency count was made, e.g., of the responses of the 400 Boston males to each of the 80 stimulus words of the LLS. Thus it was established, for example, 169 of the 400

Boston males responded to the word hammer with the word nail. Percentages were then determined for each response to each stimulus word, and distributions of percentages were reduced by taking the square root of each. This procedure aligned all the responses of the normative sample on a descending scale, according to the amount of agreement among the normative group in choosing a particular response. The resultant psychological continuum, from high to low, based upon agreement of choices among the normative subjects, is the same kind of continuum involved in previous studies on communality of association response. When the mean root frequency and the standard deviation for each stimulus - response distribution had been calculated, standard scores were computed, with an arbitrary mean set at twenty and a standard deviation at ten. This method was thought to be justified because the distributions for all the separate stimuli were of the same shape and were skewed in the same direction. It had the advantages of an equal-interval scale and the independence of means and variances.

"The final goal of scoring a test of communality must be to secure a single score for each subject for all his responses, a total score which may be taken as an indicator of the success that he has on repeated trials in finding common responses. Consequently, after the procedure described above had been carried out, test booklets of the 400 individuals in each of the four categories of subjects, namely the Boston men and women and the Chicago men and women, were scored according to the standard score scales determined from the frequency counts. Herr, V.V. (1958)."

The author of the present work had assisted Father Snider in gathering data, determining which words from the Kent-Rosanoff list were most useful and he can testify that it was a monumental task. It was at the suggestion of Fathers Herr and Snider that a further study of the possibilities inherent in the LLS, was undertaken as a piece of research at the Master's level. Since

the work just mentioned is a part of the procedure leading to the present research a description of it follows.

In September 1954, the LLS was administered to a group of college freshmen, by the author of this dissertation. The groups included men and women who were entering their first year of college. The tests were administered after the students had already taken entrance examinations and had been notified that they were admitted to college. In the general orientation meetings the students were reminded that they had been accepted but that they would have to take a series of personality tests as a part of an on-going research. They were told that these latter tests would have no bearing upon their status as students at the University. The printed instructions on the cover of the LLS booklets were read to the groups, and they were asked to re-read them carefully and to state whether or not the instructions were clear. With this manner of introducing the test, it was felt that the participants would all have the same motivation, and that they would have as uniform an understanding of the instructions as would be possible under the circumstances. Care was taken to point out that the good of science would require the fullest possible cooperation in the tests.

The booklets of all of the students who were not from the Chicago area were separated from the Chicagoans, since it was possible that geographical differences might exist in the responses. Then there were selected an equal number of booklets for the men and for the women. There were several booklets in each group which had to be rejected later on, because of failure to fill in all the answers. However the total number that were usable for the men was 96 and for the women 104.

The principal purpose of the research was to discover whether the LLS might be used as an instrument for predicting college achievement. Secondary purposes were: 1. to learn whether the LLS scores correlated with general intelligence; 2. whether the LLS will distinguish between students having high linguistic skills and those with low, and those with high performance scores from those with low; and 3. whether high achievers will score differently from low achievers on the LLS.

Since all the Freshmen had taken the American Council on Education Psychological Examination, these data were already available for the present study. Thus there was sought the degree of relationship between the LLS and the L score, between LLS and the Q score, and between LLS and the total score on the ACE.

In August of 1955, when all grades were in the dean's office and transferred to the researcher's files, each student's class standing was computed in grade points. At the same time their quantitative (Q) linguistic (L) and total scores on the ACE were secured. The LLS scores of both the men and the women were correlated with the ACE total, with the ACE Q. and with the ACE L scores. No significant correlations were found. The primary conclusion was, therefore, that the LLS, was not an indicator of the achievement of college students, and also that it was not likely to be affected by the variations in college achievement.

The hypothesis was then made that a discrepancy between the linguistic and the performance scores on the ACE might be related to scores on the LLS. In a previous study this discrepancy score had been thought to be related to personality variables, and particularly to word-association test scores. But

no significant correlations were found in this part of the study either.

Finally it was hypothesized that if the LLS is an indicator of empathy, or some form of personal adjustment, in this case adjustment to college life, then the discrepancy between general mental ability and actual college achievement might be related to scores on the LLS. Father Snider (1954) had been able, in a previous study to show some relationship between personality variables and achievement scores. He found that for high-school students, high achievers were better adjusted to their school than were low achievers matched for intellectual ability. In the present study the correlations between LLS and achievement were different for the men than for the women. For the men the Pearson coefficient of correlation was .26 and this was significant at the .01 level of confidence. This means that the higher the college achievement, the better the communality scores for men. This might be taken to indicate that the men who got along better in college and satisfied their teachers by getting good grades were also inclined to be those who shared their thought with those of other persons, and hence were able to get better communality scores on the LLS. It was previously pointed out that the communality scores did not in any way correlate with the I.Q.scores, a finding which had been verified by Herr and others repeatedly.

None of the correlations for the women, however, were significant. This was the first time in our investigation that the differences between men and women became conspicuous. We shall see the same trend occurring very often in the subsequent study.

This was a brief sketch of the preliminary data that were gathered, and which led to the design of the present experiment. As mentioned above, the

next step would be to wait another three years after the comparison had been made between entrance LLS scores and first year achievement scores, and then give the LLS over again, along with several other tests of interests and values. The deans and their assistants and secretaries proved indispensable for the gathering of the data for the later study. There had to be a search for the remaining students, after the lapse of three years, and an attempt to contact the seniors in the spring of their senior year, so that they might be induced to continue to give their cooperation in the research project which they had been responsible for beginning in 1954. As will be seen there was much more than a 50% drop-out, and this takes into account not only those who did not finish college but also those who could not be contacted, for some reason or another.

The only motivation used to secure cooperation of the students was the comment that as one of their last acts for their alma mater, they would be asked to aid in finishing the research project they had begun. They were asked to cooperate out of loyalty to their school, and no monetary remuneration was offered. Special appointments were made and the battery of tests was administered to small groups of four or five, or in most cases, to each individual separately.

As in the first administration in 1954, the booklet instructions were emphasized, and time was allowed for asking questions if these were not clearly understood. Some of the students said they vaguely remembered having taken such an association test upon entrance four years previously, but they all denied remembering just what response they might have given to any particular one of the stimulus words. This was not surprising, since Herr reports that

even on a retest taken four months after the first test, subjects were never sure which response they had given previously to any one of the stimuli; the very nature of the stimuli predisposed to this forgetting, since there is such a conglomeration of ideas represented among them; for example the list contains "stomach," "eating," "bread" and "butter" among others. A count of the total number of responses made and reported by Logsdon (1960) showed that the most frequent of all responses was "food" and that it was made in response to any one or all of the above-mentioned four stimuli. The stimuli themselves, then, tend to permit easy forgetting of the responses made.

Upon completion of the LLS tests, the students were given the Scale of Values, and the Kuder Preference Record in that order. The instructions of these tests are self-explanatory, and the testing time was under an hour. One other little motivating factor was used in order to assure seriousness in taking the tests. Students were told that the scores on the two latter tests could be very helpful in deciding upon one's vocation in life. They were promised that they could return for their actual scores after the booklets had been scored. About one third of the students indicated that they would like to receive their scores and did return to discuss them.

Having acquired the necessary data, it was now possible to determine the answers to some of the questions raised in this investigation. Thus it was possible to get the test retest reliability coefficient for the LLS after the lapse of a period of time as long as four years. Also it was possible to study the relationship between the LLS scores on the one hand and the values measured on the Allport-Vernon-Lindzey scale, and between the LLS and the Kuder preference on the other. Further, having collected the data on the frequencies

and kinds of extra-curricular activity in which the students engaged while in college, it was also possible to study the relationships existing between frequency of such activities on the one hand and good or poor communality scores on the LLS on the other; between frequency of such activities and high or low scores on the ten items of the Kuder preference record, and high and low scores on the Allport-Vernon-Lindzey scale of values. As will be noted later, special kinds of statistics will be used in testing the differences between the small groups that were used in this study, the distribution-free non-parametric statistics recommended by R. A. Fisher (1956).

CHAPTER IV

DISCUSSION OF RESULTS

The answer to the most important question in this research, namely that concerning the reliability of the LLS after a lapse of four years, was in the affirmative. The correlation coefficient between the 1954 and the 1958 LLS scores for the entire group of 17 men and 23 women was .68 whereas that for the men alone was .82, and for the women alone it was .62; it is apparent that the long-term reliability is no different than the short-term reliability studied by Herr in 1954. This is true for our present sample at least, in which the time is four years between test and retest, and the subjects were male and female college students.

In testing the hypotheses that there should be a relationship between the LLS and certain interest levels, as indicated by the Kuder Preference Record; and between the LLS and certain values as shown by the Allport-Vernon-Lindsey Scale, the rank-difference formula was used. This was required due to the skewed nature of the distributions of Kuder and Allport-Vernon-Lindsey scores. These correlations are found in Tables I through IV, on the following pages.

From Table I it appears that none of the correlations is significant even at the .05 level of confidence. A value of .48 is required for significance at the .05 level of confidence. There appears to be a tendency for high political and high economic values to be accompanied by good communality scores on the LLS, both in the 1954 and in the 1958 samples of the LLS. There appears

TABLE I

Rank Difference Correlations between Allport-Vernon-Lindzey Scale of Values and the Loyola Language Study for 17 Men, Tested in 1954, Retested in 1958

<u>LLS 1954</u>	<u>LLS 1958</u>
Religion..... .10	-.02
Political..... .35	.23
Social..... .16	.00
Aesthetic..... -.35	-.32
Economic..... .39	.33
Theoretical..... -.07	-.16

also some tendency for low aesthetic values to be accompanied by good communality scores in 1954 and in 1958. It seems logical to suppose that the men who are high in political and economic values would also be successful in judging the kind of associations other people would make on the LLS.

In collaboration with the author in 1958 a study was done by Herr and Nicolay (1960) on 78 college men, in order to check the results of this study.* Their economic scores correlated with their LLS scores to the extent of .24 which was significant at the .05 level of confidence. The relationship between political values and LLS and the opposite one between aesthetic values and the LLS were not confirmed by the Herr-Nicolay study.

*Data for this study are on file in the Lake Shore Campus Psychology Laboratory.

TABLE II

Rank Difference Correlations between Allport-Vernon-Lindzey Scale of Values and the Loyola Language Study for 20 Women
Tested in 1954, Retested in 1958

	<u>LLS 1954</u>	<u>LLS 1958</u>
Religion.....	-.24	.02
Political.....	.12	-.03
Social.....	-.07	.23
Aesthetic.....	.31	.10
Economic.....	.05	.02
Theoretical.....	-.08	-.02

From this table it appears again that none of the correlations is significant even at the .05 level of confidence. The value would have to reach .44 for significance at this level. Moreover there is not as much consistency in the correlations here as there was for the men. There is greater variance between the correlations with the LLS in 1954 and those in 1958.

On the other hand in a repeat study done in 1958 by Herr and Nicolay, using 30 college women, there are some correlations which were expected. In the Herr-Nicolay sample, religious values correlate negatively with the LLS to the extent of -.27, which is near the 1954 Stewart figure; and the aesthetic values correlate .26 which is also near the 1954 Stewart figure. On the whole the relationship between LLS scores and the Scale of Values scores for women seems to be a rather unstable one; few if any predictions could be made, even if one were to combine the probabilities derived from Stewart's sample with those from the Herr-Nicolay sample.

TABLE III

Rank Difference Correlations between Kuder Preference Record Interest Scores and the Loyola Language Study Scores for 17 Men Tested in 1954, Retested in 1958

	<u>LLS 1954</u>	<u>LLS 1958</u>
Mechanical.....	.17	.37
Computational.....	.10	.13
Scientific.....	.00	.10
Persuasion.....	.34	.30
Artistic.....	-.27	-.28
Literary.....	-.16	-.12
Music.....	-.09	.03
Social Service.....	-.08	.02
Clerical.....	-.01	.10

From this table it seems that men with high persuasive interests tend to have good communality scores although the correlation does not reach the required value of .48; those with low artistic interests tend to have good communality scores. These results corroborate, to a certain extent, the results of the study with the Allport-Vernon-Lindzey Scale of Values. Men with low artistic interests as shown on Table III, and those with low aesthetic values, as shown by Table I do better at sensing the commonest associations made by their fellow men on the LLS. Moreover, if men are persuasive, they tend to sense their fellow man's associations better than otherwise.

In an attempt to cross-validate these findings, the author along with Herr and Grant in 1959, using a sample of 22 college boys, found the correlations with persuasive interests to be .26. When this value was taken in

conjunction with those for the 1954 and 1958 samples by combining the probabilities, we have good evidence that the relationships are not due to chance. The significance goes beyond the .05 level of confidence (Lindquist, E. F., 1953). However, the negative correlation between artistic interests and the LLS, for the Herr-Grant sample, became zero. Consequently one could not make predictions with regard to the artists.

TABLE IV

Rank Difference Correlations between Kuder Preference Record Interest Scores and the LLS Scores for 20 Women Tested in 1954 and Re-tested in 1958

<u>LLS 1954</u>	<u>LLS 1958</u>	
Mechanical.....	-.35	-.22
Computational.....	.04	.27
Scientific.....	-.28	..24
Persuasive.....	.05	-.08
Artistic.....	.23	.14
Literary.....	.48	.37
Musical.....	.22	.08
Social Service.....	-.05	.13
Clerical.....	.17	.00

Here there is a correlation which is significant at the .05 level of confidence, for literary interests correlated with the 1954 LLS. The women with low scientific and mechanical interests tend to do well on the LLS. The correlations seem to be opposite to those for men. In an attempt to cross-validate the same, Herr and Grant (1960) using a sample of 28 women and a different college, found a positive correlation with artistic interests of .22, and by combining this with the figure from the Stewart sample, we have a

significance beyond the .05 level of confidence. None of the other correlations was corroborated. Thus again the women are more inconsistent than the men. Nevertheless it is perhaps safe to predict with some degree of confidence that highly artistic women sense the associations of other women better than their colleagues who possess these interests to a less degree.

Since in the 1954 study data had been secured from which it could be determined which students showed clear behavioral evidences of literary interests, it was decided to dichotomize all of the women into two groups. One group will be called the literary group and the other the non-literary group. The author was prompted to follow this question further because of the correlation just mentioned. He wished to ascertain whether or not the literary group would score better on the LIS than the non-literary group. He was able to make the dichotomy on the basis of the fact that he had kept a record on the number of activities in which all the students participated while in college. Examples of over-all activities might be belonging to fraternities and sororities, participation in extracurricular activities and the like. Membership in study clubs, being on editorial staffs of school papers, belonging to dramatic societies were examples of literary activities.

When the lists of these activities had been compiled, it became evident that the women tended to belong to quite different types of outside activities than did the men. Whereas women joined "study clubs," sororities, literary and dramatic societies rather more frequently than men, and whereas men belonged to athletic and recreational and scientific societies more often than women, it became necessary sometimes to use different sets of categories for men than for women.

The women were first therefore categorized into literary and non-literary, on the basis of their amount of participation in literary activities alone. The men and the women were categorized subsequently into both the following groupings: those with no such activity versus those with some; and those with no such activity or only ONE OF THEM versus those with MORE THAN ONE activity. Then each of these pairs of groups was tested for their communality of thought as measured by the LLS.

The tabulation showed that 18 women could be put in the category of literary, and 42 in the non-literary group. The LLS scores for the 18 had a mean of 505.69 with S.D. of 66.44 whereas the LLS scores for the 42 had a mean of 496.83 with S.D. of 46.21. The difference between these two means is smaller than the standard error of the difference.

Next there was made a comparison of the LLS scores of the women according to the other above-mentioned method of groupings.

There were only 6 women who had no activities at all, and so this category was omitted entirely. However there were 29 with either one or no outside activity (Group II). These had a mean LLS score of 511.33 with an S.D. of 69.53. The women with more than one activity (Group I), numbered 31 and their mean LLS score was 495.16 with an S.D. of 51.09. These two means also differ so slightly that the difference could well be attributed to chance.

When the men are considered there is a slightly different story. Of the total of 53 men whose records were kept, there were 19 with no activities that could be found. Their mean LLS was 502.89 with an S.D. of 63.07 whereas the mean of the 34 men with one or more activities averaged 490.52 with an S.D. of 73.23; again the mean difference is slight. Following the division which was

used with women the investigator now divided the men into those with ONE OR NO activities (Group II) versus those with MORE THAN ONE (Group I). In Group II there were 38 and their mean LLS score was 505.39 with S.D. of 64.26. In Group I there were 15 and their mean LLS score was 468.47 with an S.D. of 54.53. Here is a mean difference that is significant by the t test at the .01 level of confidence. Those gentlemen who join more activities while at school show more communality of thought on the LLS. Thus on the whole the search for behavioral correlates of the LLS scores has not been very successful with the men and women chosen for this study.

It may be helpful to recall here that the rationale of the present investigation was primarily to test for the stability of the LLS scores over a rather extended period of time. If such stability of the LLS proved to be the case, then it was thought that the rather abiding interests and values of the students, measured in their last year of college, might bear some relationship to either of the two LLS scores or to both. Up to this point some few rather tenuous conclusions have been justified about the relationship between interests and communality of thought on the one hand, and between values and communality on the other.

It is our purpose now to push this investigation somewhat further and to compare the mean Kuder Preference scores of the highly active groups with mean scores of the less active groups, both for the men and for the women. The thinking grew out of the observed facts mentioned above, namely that a tendency to engage in much group activity seems related, at least in one of the group-pairs studied, to communality of thought. Moreover, as seen in Tables I through IV, certain of the Kuder Preferences, and certain of the

Values appear related to communality of thought. Therefore it could be suspected that tendencies to activity in college might be used to discriminate certain preferences for occupations or job preferences, from other job preferences; and that the same activity tendencies might be used to discriminate certain Value levels from others. In a word, the question is asked whether or not the highly active college students show different basic preferences and interests, certain systems of value that are different from those of the less active students.

If the highly active group, then, does have these different Preferences and Values, these differences might be looked upon as grounds or bases for the differences that have been discovered on the LIS test of communality of thought between high and low activity persons. They might also be grounds for the trends in the correlations, between the communality scores and the Preferences Scores and Value levels.

To become more specific, the high activities group would be expected to be more like persons who score high on the Kuder persuasive scale, and on the social service scale, and less similar to those who score high on the science and computational scale.

Again highly active persons might be expected to lean more strongly toward choosing social and political systems of value; not so much toward choosing aesthetic value systems. And just these differences in value systems might be at the basis for the discovered differences in communality of thought between the highly active and the less active groups, as well as for the few significant and near significant correlations between the communality scores and the preferences and values.

However, since the women in our study usually though not always had opposite tendencies to those of the men, as shown by their correlations and their mean differences between highly active and less active persons on the LLS, it is predicted that none of the differences in mean Allport value scores, and none of the differences in Kuder preferences for the women's active and inactive groups will be significant. One lone exception might be traceable to the high positive correlation between women's literary interests and their communality scores.

In making the comparisons between the highly active groups and the less active groups, the same dichotomy will be kept as before, in that those men and women who had NONE or ONLY ONE activity will be called the inactive group or Group II, and those with MORE THAN ONE will be called the Active Group or Group I. For the men, there will be 38 in the inactive Group and 15 in the active Group; and for the women there will be 29 in the inactive Group and 31 in the active Group.

In the following pages the mean Kuder scores for the two groups, called Group I and Group II, will be found; Table V shows the two sets of Kuder scores for the men; Table VI shows the two sets of scores for Values for the men; Table VII shows the same for the women as Table V; Table VIII shows the same for the women as Table VI. Tables IX, X and XI will be explained later.

The means and mean differences are listed without standard deviations. This had to be done owing to the fact that the interest and value scores are known to have skewed distributions. Instead, the Exact Probabilities Test of

Fisher will be applied to those differences which appear to approach significance. This test makes no assumptions about the nature of the distributions, such as would be necessary if other parametric statistical procedures were used. The Fisher Test is receiving wide use today in education and in psychology

TABLE V

Mean Scores for Men, and Differences between Group I and Group II on the Kuder Preference Record. Group I are the High-Activities and Group II are the Low-Activities Persons

	Group I	Group II	Differences
1. Mechanical	21.60	24.38	2.78
2. Computational	19.40	26.00	6.60
3. Scientific	22.80	31.31	8.51
4. Persuasive	60.00	37.54	22.46 *
5. Artistic	20.60	29.31	8.71
6. Literary	34.40	29.23	5.17
7. Musical	19.00	22.38	3.38
8. Social Service	46.20	37.46	8.74
9. Clerical	42.60	43.23	.63

There is here only one rather large difference which is marked with an asterisk. The high activities men have the larger mean. Men who join more groups have more similarity to the persuasive and social service type of person. The active type, however, appears to have lower artistic and scientific interests. The findings, regarding persuasive and artistic interests are consistent with the correlations listed in Table III, but the finding on social service is not.

TABLE VI

Scores on the Allport-Vernon-Lindzey Scale of Values made by Men. High Activities Persons are Group I and Low Activities are Group II

	Group I N 15	Group II N 38	Differences
1. Theoretical	36.62	47.62 *	* 11.00
2. Economic	36.88	32.33	4.55
3. Aesthetic	35.12	44.87 *	* 9.75
4. Social	34.50	33.33	1.17
5. Political	46.00	41.73	4.27
6. Religious	43.12	47.66	4.54

Here the two larger differences are those with theoretical and aesthetic values. In both these the low activities group of men scores higher.

Consulting Table I we see that the men's aesthetic and theoretical interests also were negatively correlated with their communality scores, in such wise that the better the communality scores the lower the aesthetic and theoretical values. Perhaps the lack of group activities could be somehow related to heightening of aesthetic and theoretical interests at least for men.

There is no difference between means of the low and high activities groups which is greater than 8.57 and there is only one correlation between the women's interests and the LLS that was near significance. It was their literary interests. Yet the literary interests are no different for the active and the inactive Groups. There is no consistency between the findings of the LLS and those of the Preference Records for active and inactive women.

TABLE VII

Scores on the Kuder Preference Record made by Women. High Activities Persons are Group I, and Low Activities are Group II

	Group I N 31	Group II N 29	Differences
1. Mechanical	26.00	21.00	5.00
2. Computational	19.85	24.00	4.15
3. Scientific	31.85	39.78	7.93
4. Persuasive	34.46	25.89	8.57
5. Artistic	27.85	28.67	.82
6. Literary	25.69	27.22	1.53
7. Musical	14.46	18.56	4.10
8. Social Service	52.15	54.67	2.52
9. Clerical	40.38	42.44	2.06

TABLE VIII

Scores on the Allport-Vernon-Lindzey Scale of Values made by Women. High Activities are Group I, Low Activities Group II

	Group I N 31	Group II N 29	Differences
1. Theoretical	39.20	41.44	2.24
2. Economic	33.93	29.22	4.71
3. Aesthetic	41.07	42.22	1.15
4. Social	39.47	41.00	1.53
5. Political	37.47	34.11	3.36
6. Religious	49.40	55.33	5.93

For women there are no value differences greater than 5.93 between the high and low activity persons. There were also no consistent correlations

between values and communality of thought as shown in Table II for women. However, it was seen there that the poorer the communality scores, the higher the religious values, though the correlation was not significant. Here we see that the low-activity group has higher values along religious lines. It seems consistent to say that poor communality (empathy), low social or group activity, and HIGH religious (ascetical) values could go together.

As stated above, the means for the active and the non-active groups ought not be tested for significance of differences by parametric tests. Now there will be applied, however, in order to get the most information out of the data, the Fisher (1956) Exact Probabilities Test. The results for men are shown in Tables IX and X; those for women in Table XI. The differences in the case of women on the Scale of values are virtually non-existent and hence they are not shown here. Those in the case of men show one significant difference in line with what was predicted, namely that the active group has higher persuasive interests, Table IX. Another difference approaches significance, and that is between the active and non-active and on both the theoretical and aesthetic values, Table X. Those who are not over active have higher values along these two lines of human endeavor, and this seems consistent with what one would expect.

The other interests were not even close to significance, but we can say that in a group of men, those with high communality scores will be more active and that they will be more persuasive than the less active persons. This corroborates the real difference in this area reported on page 39.

TABLE IX

Fisher's Exact Probabilities Test, to Discover the Significance
of the Differences between the Active and Non-Active Scores
of Men on the Kuder Preference Record

	Median		NA	A	
Computation	23 $\frac{1}{2}$	Above Median	7	2	Not significant
		Below Median	6	3	
Scientific	26 $\frac{1}{2}$	Above Median	8	1	Not significant
		Below Median	5	4	
Persuasive	43	Above Median	4	5	Significant at the .025 level as expected. A's have higher persuasive interests
		Below Median	9	0	
Artistic	25	Above Median	8	1	Not significant
		Below Median	5	4	
Social Service	37	Above Median	6	3	Not significant This contrary to expectation
		Below Median	7	2	

NA means non-active

A means active

TABLE X

Fisher's Exact Probabilities Test to Discover Significance of Difference
between Active and Non-Active Men on Two Scale of Value Scores
(Theoretical and Aesthetic) on the Allport-Vernon-Lindzey
Scale of Values

	Median		NA	A	
Theoretical	38	Above Median	11	2	.06 level of confidence
		Below Median	6	6	
Aesthetic	41	Above Median	11	2	.06 level of confidence
		Below Median	6	6	

NA means non-active A means active

There is a definite trend shown for inactive men to show higher theoretical and aesthetic values than the active group. This same strong tendency was noted in Table VI, page 44.

TABLE XI

Fisher's Exact Probabilities Test, to Discover the Significance
of the Differences between Active and Non-Active
Women's Scores on the Kuder Preference
Record (Scientific and Persuasive)

	Median		NA	A
Scientific	32	Above Median	6	5
		Below Median	3	8
Persuasive	27 $\frac{1}{2}$	Above Median	3	8
		Below Median	6	5

NA means non-active A means active

While neither of these is near significance, still it may be said that non-active women have more scientific interests, and active women more persuasive interests.

These are the same tendencies which we had expected and found for the men, though not so pronounced.

CHAPTER V

SUMMARY AND CONCLUSIONS

1. The four-year test - retest reliability of the ILS is .68 for the men and women combined. It is higher for men than for the women. The four month reliability is of the same order for a sample of over 200 men and women. In this latter case the coefficient was slightly higher also for the men than for the women.
2. The results of the present study verify some of Halpern's findings and predictions.
3. There is a tendency for men with low aesthetic values to score well on communality of thought; but for men of high political and economic values to score well on the ILS, the larger samples having confirmed only the relationship between the ILS and economic values.
4. There were no correlations between the ILS scores and the values for women, neither in the present small sample nor in the larger samples.
5. The present study showed a tendency for men with high persuasive interests and low artistic interests to do well on the ILS. A confirmatory study had shown the same tendency for persuasiveness, but not for artistic interests.
6. For the women of the present study good ILS scores are positively related to literary interests, whereas they are related negatively, but not significantly, to mechanical and scientific interests. The follow-up failed to corroborate these findings, yet the tendency appeared for artistic

interests to go with good communality scores.

To summarize thus far: for men, high persuasive interests and economic values tend to be related to good IIS scores; for women no very strong relationship exists save that between literary-artistic interests and good communality scores.

The tests for the significance of the difference between the active and the non-active group confirm only the following hypotheses:

- A. For men the active ones have significantly better Kuder persuasive scores than do non-active ones. On the scale of values, there exists only a slight tendency for the inactive group to score better on theoretical and aesthetic interests than the active group.
- B. For women: not one of the hypotheses was confirmed. A clear tendency existed for the active group to score better on persuasive interests and more poorly on scientific interests; also, those women who do not join sororities in college tend to have slightly higher religious values. Finally, although it may seem that the results of this study have been mostly negative, still they are very decisive and useful in their application. That they are of value stems from the fact that they may suggest the idea to others, either NOT to try any longer to find the correlates between personality functioning and communality of thought as measured by the IIS, or precisely to try to cross validate the above findings, especially the clear differences which appeared between men and women.

The single crucial positive result is of course, the presentation of

positive evidence for the four year reliability of the LLS. And, as a matter of fact the LLS, alone or in combination with other tests, is actually being used tentatively as part of a screening device for selection of persons best suited for clerical and religious life. It has possibilities, as we see from the findings of this research, in helping to discriminate men with high persuasive interests from those with low persuasive interests.

Thus, it would seem that Halpern's findings in 1957 have been partially verified in this investigation. It will be recalled that he found some relationship between empathy and values. More definitely, he found empathy in women positively related to social values but negatively with aesthetic values. In our study the LLS scores, assuming that they do measure some sort of empathy, are positively related to aesthetic values in women, but negatively, to religious values, whereas for the men empathy is negatively related to aesthetic values, at least in the one sample of 17 men. For the men of both the Stewart and the Herr-Nicolay samples, economic values were positively related to empathy as measured by the LLS.

As is apparent there is need for much more research in this whole area of empathy, which is of such supreme importance for studies of human interpersonal relationships. The directions which the further research might profitably take are: 1. toward larger samples for ascertaining the test-retest reliability of the LLS; 2. studies of the relationship between communality of thought on the LLS and other measures of empathy; 3. a closer look into the possibility mentioned by Jenkins in 1959, of finding word-association scores related to "social sensitivity"; 4. pursuing further the extent of the sex differences which seemed so prevalent in this study. To cite one example only:

men high in persuasive interests, and in economic values do better on the ILS than men low in these interests; conversely men low in aesthetic values and artistic interests tend to do better on the ILS than do those who score high. Neither of these relationships hold for the women studied in this research.

BIBLIOGRAPHY

- Allport, G. W. *Personality: a psychological interpretation*, N.Y.: Holt & Co., 1937.
- Carter, H. D. and Sorenson, M. L. Twin resemblance in community of free association responses. *J. Psychol.*, 1940, 9, 237-246.
- Clousing, Josephine. The relation between personality trait ratings and free association responses. Unpublished Master's thesis. Univ. of Minnesota, 1927.
- Conrad, H. and Harris D. The free association method and the measurement of adult intelligence. *Univ. Calif. Publ. Psychol.*, 5, 1931, 1-45.
- Cronbach, Lee J. *Essentials of psychological testing*, Harper & Brothers, pp. 348-49, 1949.
- Crosby, R. C. and Winsor, A. L. The validity of student's estimates of their interests. *J. appl. Psychol.*, 1941, 25, 408-414.
- Del Vecchio, A. The discriminatory power of the Loyola Language Study in schizophrenia. Unpublished Ph.D. Dissertation, Loyola University, Chicago, 1957.
- Dinello, Frank. Loyola Language Study and its relation to various professional groups. Unpublished Master's thesis, Loyola University, Chicago, 1958.
- Fisher, R. A. Cited in Siegel, Sidney. *Non-parametric statistics for behavioral sciences*, pp. 96-104. N.Y., McGraw-Hill, 1956.
- Goodenough, Florence L. The use of free association in the objective measurement of personality. In G. McNemar & M. A. Merrill (Eds.), *Studies in personality*. New York; McGraw-Hill, 1942, pp. 87-103.
- Grant, G. Unpublished Research. Loyola University, 1960.
- Halpern, H. M. Predictive empathy and the study of values. *J. Consult. Psychol.*, 1957, 21, 104.
- Harris D. and Conrad H. The free association method and the measurement of adult intelligence. *Univ. Calif. Publ. Psychol.*, 5, 1931, 1-45.
- Herr, V. V. The Loyola Language Study. *J. Clin. Psychol.*, Vol. XIII, 1958, No. 3, 258-262.

- Jenkins, James J. & Russell, Wallace A. Systematic changes in word association norms. *J. Abn. & Soc. Psychol.*, 1960, 60, 293-303.
- Jenkins, James J. Effect on word-association of the set to give popular responses. *Psych. Reports*, 1959, 5, 94.
- Jung, Carl. The association method. *Amer. J. Psychol.*, 21, 1910, 219-269.
- Keene, C. Mansel. Commnality of response on a word association test: a study of standardization procedures and an attempt to forecast moderate emotional adjustment. Unpublished doctoral dissertation, Stanford Univ., 1951.
- Kent, G. H. & Rosanoff, A. J. A study of association in insanity. *Amer. J. Insanity*, 1910, 67, 37-96, 317-390.
- Leads, B. S. and Tresselt, M. E. The Kent-Rosanoff word association: I. New frequencies for ages eighteen-twenty one and a comparison with Kent-Rosanoff frequencies. *J. Genet. Psychol.*, 1955, 87, 145-148.
- Lindquist, E. F. Design and analysis of experiments in psychology and education. Houghton Mifflin, Boston, 1953.
- Logsdon, Laura. Unpublished M.A. Thesis, Loyola University, 1960.
- Malamud, B. I. Value of the Maller controlled association test as a screening device. *J. Gen. Psychol.*, 1946, 21, 37-43.
- McCall, W. C. and Traxler, A. E. Some data on the Kuder Preference Record. *Educ. psychol. Measmt.*, 1941, 1, 253-268.
- Nicolay, Robert. Unpublished research project, Loyola University, 1958-1959.
- O'Connor, J. Born that way, Baltimore: Williams and Wilkins, 1928.
- Rosanoff, Aaron J., (Ed.) Manual of Psychiatry, Sixth edition. John Wiley & Sons, Inc., N.Y., 1927.
- Rosanoff, Isabel R. and Rosanoff, A. J. A study of association in children. *Psychol. Rev.*, 1913, 20, 43-89.
- Russell, Wallace A. and Jenkins, James J. Systematic changes in word association norms. *J. Abn. & Soc. Psychol.*, 1960, 60, 293-303.
- Snider, L. B. In the human person (eds. Arnold, M. B. and Gasson, J.) A research method validating self-determination. Chapter 7, p. 222. Ronald Press, 1954.
- Sorenson, M. L. and Carter, H. D. Twin resemblance in community of free association responses. *J. Psychol.*, 1940, 9, 237-246.

- Spranger, Ed. *Psychologie des Jugendalters*. Leipzig, Quelle and Meyer, 1924.
- Stanek, Richard. An investigation of the influence of age, sex and education on responses to a semi-controlled association test. Unpublished Ph.D. Dissertation, Loyola University, Chicago, 1956.
- Stewart, J. V. P. The Loyola Language Study and college achievement. Unpublished Master's thesis. Loyola University, Chicago, 1956.
- Tresselt, M. E. and Leeds, D. S. The Kent-Rosanoff word association: I. New frequencies for ages eighteen-twenty one and a comparison with Kent-Rosanoff frequencies. *J. Genet. Psychol.*, 1955, 87, 145-148.
- Traxler, A. E. and McCall, W. C. Some data on the Kuder Preference Record. *Educ. psychol. Measmt.*, 1941, 1, 253-268.
- Triggs, Frances O. A study of the relationship of measured interest to measured mechanical aptitude, personality, and vocabulary. *Amer. Psychologist*, 1947, 2, 296.
- Winsor, A. L. and Crosby, R. C. The validity of student's estimates of their interests. *J. appl. Psychol.*, 1941, 25, 408-414.

APPROVAL SHEET

The dissertation submitted by John V. P. Stewart has been read and approved by five members of the Department of Psychology.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Jan 15, 1963
Date

Wanda L. Hester
Signature of Adviser