A Study of the Relationship between Communality of Thought on the Loyola Language Study and Empathic Ability on Kerr's Empathy Test

Mary Lucinia Szpak

Loyola University Chicago

Recommended Citation
http://ecommons.luc.edu/luc_theses/2214
A STUDY OF THE RELATIONSHIP BETWEEN COMMUNALITY OF THOUGHT
ON THE LOYOLA LANGUAGE STUDY AND EMPATHIC ABILITY
ON KERR'S EMPATHY TEST

by
Sister Mary Lucinia Szpak, C.S.S.F.

A Thesis Submitted to the Faculty of the Graduate School
of Loyola University in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

February
1967
LIFE

Sister Mary Lucinia, C.S.S.F., was born in Chicago, Illinois, on June 25, 1923.

She was graduated from Good Counsel High School, Chicago, Illinois, and holds a Bachelor of Science Degree in Education and a Master of Education Degree from Loyola University.

From 1948 to 1966 the author was engaged as a teacher on the elementary school level. Twelve years of this time, from 1954 to 1966, were spent at one of the centers for deaf children in the Archdiocese of Chicago.

As lecturer in the Institute of Rehabilitation, the writer taught courses in reading and language for the deaf at DePaul University from 1962 to 1964.

Presently, the writer is pursuing post-graduate work at Loyola University.
ACKNOWLEDGEMENTS

The author wishes to express her appreciation to Reverend Vincent V. Herr, S.J., for his invaluable suggestions, continued guidance, and patient understanding throughout the study.

The writer also feels deep indebtedness to Mother Mary Innocenta, C.S.S.F., and the Sisters of her Community for their interest in the study, encouragement while it was in progress, and assistance in bringing it to completion.

Special grateful acknowledgement is extended to Sister Mary Ursuline, C.S.S.F., Mistress of the Juniorate; to Sister Mary Laudiose, C.S.S.F., Mistress of the Novitiate; and to Sister Mary Avila, C.S.S.F., Mistress of the Postulancy, for making available the subjects for this study.

Heartfelt thanks are also extended to the Postulants, Novices, and Junior Sisters who participated in this study.
TABLE OF CONTENTS

LIFE OF AUTHOR ......................................................... i
ACKNOWLEDGEMENTS ................................................... iii
LIST OF TABLES ........................................................... iv
LIST OF ILLUSTRATIONS ................................................... v

Chapter

I. INTRODUCTION ....................................................... 1
II. REVIEW OF RELATED LITERATURE .................................. 6
III. METHODOLOGY-DATA COLLECTION .................................. 27
IV. DISCUSSION OF RESULTS ............................................. 48
V. CONCLUSION ............................................................ 56

BIBLIOGRAPHY ............................................................. 60

APPENDIXES

I. LOYOLA LANGUAGE STUDY ............................................. 65
II. KERR'S EMPATHY TEST ................................................ 66
III. THE ADJUSTMENT RATING SCALE ..................................... 67
IV. NORMS ON THE EMPATHY TEST ....................................... 68
V. HISTOGRAM OF LOYOLA LANGUAGE STUDY SCORES ................. 69
VI. HISTOGRAM OF EMPATHY TEST SCORES ............................... 70
VII. HISTOGRAM OF ADJUSTMENT RATING SCORES ..................... 71
LIST OF TABLES

Table | Page
---|---
1. Tabulation of Subjects According to Membership Groups, Mean Group Age, and Time in Religion | 27
2. The Loyola Language Study. Discriminating Words for Chicago Males (M) and Females (F) | 36
3. Number of Subjects Achieving Percentile Ranks Comparable to Norms for Liberal Arts Women on the Empathy Test | 49
4. Number and Percent of Subjects Achieving Scores at $Q_3$ (or above), $Q_2$ (or above), $Q_1$ (or above), and below $Q_1$ Based on Norms in Table 6 | 50
5. Range of Desirable and Undesirable Scores on the Loyola Language Study and the Empathy Test for 118 Subjects | 51
6. Range of Scores for Each Descriptive Category on the Adjustment Rating Scale | 52
7. Means, Sigmas, and Ranges of Scores on the Loyola Language Study, the Empathy Test, and the Adjustment Rating Scale | 52
8. A Comparison of the Frequency Count for Each Rating Category with Frequency Count of Desirable and Undesirable Scores on the Loyola Language Study and the Empathy Test | 53
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Distribution of scores of 118 subjects on the Loyola Language Study</td>
<td>69</td>
</tr>
<tr>
<td>2.</td>
<td>Distribution of scores of 118 subjects on Kerr's Empathy Test</td>
<td>70</td>
</tr>
<tr>
<td>3.</td>
<td>Distribution of scores of 118 subjects on The Adjustment Rating Scale</td>
<td>71</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

One aspect of the general area of interpersonal dynamics concerns itself with a type of human response to others that psychologists denote by the term empathy. It is a form of detached insight into the attitudes, opinions, or feelings of another person without becoming emotionally involved.

Kerr and Speroff (40) define it as "the ability to put oneself in the other person's position, establish rapport, and anticipate his reactions, feelings, and behavior." Grossman (28) refers to empathy as the ability "to perceive correctly the world from the other person's frame of reference." Dymond (20) categorizes it as a kind of role playing calling for an "imaginative transposing of oneself into the thinking, feeling, and acting of another and so structuring the world as he does." No matter what terminology is used to define it, basically the empathic response emerges as a dynamic process calling for continuous selective perception, structuring, participation, and prediction.

Empathy, however, should not be confused or equated with a similar and related concept such as sympathy, nor with the psychological mechanisms of identification and projection. Although these concepts may be corporate factors contributing to the concept of empathy, they are, nevertheless, distinguishable from it.
A little reflection reveals that the ability to empathize is of central importance in all social interaction and group integration. Its significance extends far beyond its relevance for psychiatric phenomena. Teachers, nurses, physicians, clinicians, counselors, managers, politicians, salesmen, and social workers—all utilize some degree of empathic knowledge as part of the human relations skill demanded by their occupational and professional relationships. According to Katz (39)

Anyone who takes the role of a helper in psychological and subjective concerns of other people is confronted with the complexities of empathic involvement no matter what level of depth the relationship takes. Some relationships are more emotionally charged than others, but proficiency in empathy is almost always expected of the helping person if he is a professional and responsible individual. Psychoanalysts, physicians, and counselors, for example, do not enter the same relationships with the individuals they help; but empathy is a common denominator among them.

Our everyday experiences also give evidence to the fact that we depend necessarily on our capacity to perceive and predict the behavior, thoughts, and feelings of the other person if we want to live harmoniously with him. Even the subtle interchange of love and friendship rests on skill in perception and prediction.

Having delineated the importance of the empathic response in all social interaction, the writer raises the following question regarding the concept of empathy. Is it a cause or an effect of good interpersonal relationships? Parker (63) maintains that despite the lack of adequate empirical evidence supporting the causal nature of empathy, the rationale of the concept prompts the conclusion that empathy is a cause more than an effect of good interaction with people. Therefore, it would be reasonable to assume that persons are
good leaders, warm friends, and effective in their relationships with others because they have the ability to empathize. This assumption, however, does not preclude the possibility that empathy may be enhanced by various interpersonal relationships.

If empathy is a cause of good interpersonal relations, it seems reasonable to this investigator that although it is, in a sense, an unseen phenomenon, its measurement is a sufficiently important topic in psychology to warrant attention.

The most commonly used approach to the measurement of empathy has been to require subjects to predict the responses of other persons on a rating scale or personality test. In this study, however, the instrument chosen to measure empathic ability is Kerr's Empathy Test. Based on the assumption that empathic ability may be demonstrated by predicting the behavior of the average or hypothetically average person, the Empathy Test requests the testee to predict the responses of typical individuals in three key areas: (1) aesthetic (music-types), (2) general human interests (what people read), and (3) interpersonal relations (annoying experiences). Van Zelst (67) refers to this type of empathizing as mass empathy since one must not only "feel into" or interject oneself into the structural field of another specific person, but also assume the average person's structural field.

Compatible with this connotation of the empathic process is another type of personality construct closely akin to mass empathy but referred to as communality of thought. Generally speaking, it is a person's sensitivity to the common associations shared by the rest of the population. This sensitivity recognized as an index to a person's mental health can be measured by
free or controlled association tests which according to Herr (35), "have long been regarded as being among the most important instruments for measuring personality."

One instrument purporting to measure communality of thought is the Loyola Language Study—a semi-controlled association test. The rationale of the test is based on the premise that there exists a communality of thought associations widely shared by the general population. Consequently, the individual's ability to give evidence of sharing these common associations is indicative of mental adjustment. This deduction can be made on the assumption that as a subject is deliberating a response to a stimulus word in a semi-controlled association test such as the Loyola Language Study, he calls into play higher mental processes which are often affected by emotional disturbances and distorted thinking. According to Kent and Rosanoff, who have been largely responsible for making the word association technique a diagnostic tool, idiosyncratic responses to the stimulus words are indicators of possible mental disturbances, incipient emotional complexes, or just worthy of further analysis.

The similarity in the definitions of empathic ability and communality of thought constitutes the basis for the present investigation. The assumptions underlying the Loyola Language Study and the Empathy Test prompt the investigator to show empirically whether a relationship exists between the scores on these two instruments. A correlation technique is used to ascertain this relationship. Consequently, it is hypothesized in this study that there is a positive correlation between desirable Loyola Language Study scores and desirable Empathy Test scores. The writer suspects that since both tests
demand of their respondents some sort of prediction regarding the generalized other, the responses, though of a different nature, seem to be the result of the same personality construct.

It is also hypothesized in this study that religious rated as well-adjusted by their associates are more likely to have desirable scores on the Loyola Language Study and the Empathy Test than those religious rated as poorly or moderately adjusted to community living. To test this hypothesis, the writer devised a rating scale to yield a numerical index of a subject's adjustment.
According to Allport (2) "no person can understand any other person completely because no human being shares directly the motives, thoughts, and feelings of another." Yet much of our lives is spent in trying to understand others and in wishing others would understand us better. Various theories of cognition have been propounded to explain the complexities of our perception of one another. Among them, the theory of inference holds an important place in the process of understanding. Allport maintains, however, "that for all its unquestioned merits, the theory of inference falls seriously short of adequacy" in person perception.

This inadequacy was recognized by psychologists, especially in Europe. Consequently, they suggested new concepts of interpersonal dynamics to supplement the limitations of the inference theory. Broader and more adequate accounts of the process of understanding people resulted.

The concept of Einfühlung (feeling oneself into) is one such supplementary concept introduced by Lipps at the turn of the century and translated as "empathy" by Titchener.

According to Lipps we acquire knowledge from three sources: (1) people, (2) inanimate objects, and (3) self. Arnold (3) labels these
sources of our knowledge and by means of an analogy presents a paraphrase of Lipps's theory of empathy which is formulated around the concept of Einfühlung. According to her translation, Einfühlung is "the source of our knowledge of others comparable to sensory perception which is the means by which we know inanimate objects, and introspection, by means of which we know ourselves."

As originally used, the concept referred primarily to the process of mimicry, whether it was applied to esthetics or to the understanding of people. Thus, within the framework of the concept, the contemplation of a work of art was described as involving many slight movements of the brows, eyes, trunk, and limbs which were thought to be in some way imitative of the stimulus-object. Likewise, attributing empathic elements to the facial expressions of a sympathetic audience and to the many judgments that are made about people also falls within the connotation of the concept. For example, statements like "her placidity has a calming effect," "his movements are invigorating," "his mirth is irresistible," or "his depression weighed me down" connote mimicry, although we are unaware of the extent to which this mimicry enters into the process of understanding.

The empathic response as described above seemingly fits well under the theory of inference (infant reactions excluded). However, in an attempt to preclude having empathy equated with "kinesthetic inference," Lipps emphasized the concept of "otherness" as part of the empathic process. With this emphasis, although we employ our past experience in empathizing, the process itself has exclusively objective reference. According to Lipps and paraphrased by Allport (2) "we do not perceive our own body in action but the body of the
other. There is no break between the strain, pride, sorrow, or playfulness which I feel empathically and the personality of the one I am seeking to understand."

Lipps's efforts to remove empathy from the simple realm of inference, however, did not restrain Allport from labeling empathy as a "half-way" theory. Consequently, he says that the empathic process is replete with kinesthetic inference in spite of the fact that Lipps gives precedence to the Thou in the process of perception.

Although empirical validation of the empathic response was slow in coming, psychologists and psychiatrists have recognized its practicable characteristics and have used it as an explanatory concept in many interpersonal relationships. Sullivan (63), for example, uses empathy to explain the emotional contagion or communion between child and significant adults among whom the mother holds precedence. He believes that from the very first days of life, the child somehow feels an empathized comfort or discomfort which the significant adults produce in him in accordance with their friendly or unfriendly attitude toward him. Hoskins (36) speaking of schizophrenia says that "perhaps as fundamentally characteristic as anything about the psychosis is the failure of the subject either to achieve or retain adequate breadth or depth of empathy."

The ramifications of the empathic process, however, extend far beyond the psychiatric realm. We need only to look at interpersonal situations for evidences. Lindgren (46) attributes the success of a speaker to his ability to empathize with his audience. Sensitive to the mood of the group, a good
speaker detects symptoms of restlessness and boredom in his audience and uses them as cues to change pace or topic.

Unquestionable is the importance of the empathic response in the role of a teacher whose work involves her in continuing interaction with people. From their studies in teacher-pupil interaction, Lindgren (46) and Dixon and Morse (18) concluded that good teachers are characterized by high empathic ability.

A doctor, too, benefits from his ability to empathize, for through it he gains insight into a patient's emotional state. As Lagemann (44) remarked "empathy enables the doctor to find out what kind of a person the disease has got."

The importance of the empathic process in the field of psychotherapy has been recognized by Patterson (50) who states that "therapeutic understanding is achieved by trying to see things from the client's point of view." In this connotation therapeutic understanding is related to empathy defined as role playing ability which is enhanced by a broad understanding of human nature. Sullivan once said, "We are all much more simply human than otherwise." Consequently, the basic similarities of human beings provide the essential foundation for empathy.

In a recent article, Rogers (53) discusses the three attitudinal patterns which a therapist must cultivate in order to effect constructive or therapeutic change in the client. One of the attitudinal patterns he describes as an accurate empathic understanding of the client's private world. In Roger's own words
To sense the client's inner world or private personal meaning as if it were your own, but without ever losing the "as if" quality, this is empathy, and this seems essential to therapeutic change. To sense his anger, or his fear, or his feeling of being persecuted as if it were your own, and yet without your own anger, fear, or suspicion getting bound up in it, this is the condition we are endeavoring to describe. It is this kind of highly sensitive empathy which seems essential to therapeutic change.

Other examples of the relevancy of empathy to all phases of life are legion. Lindgren (46) avers that empathy is an essential tool in furthering an understanding of oneself and others. Unless some aspect of our own self structure prevents us from empathizing, we attain higher levels of emotional maturity proportionally to our exercise of the empathic response.

In comparison with the study of other psychological phenomena, the efforts, in terms of empirical studies attempting to clarify the concept of empathy, have been practically non-existent for more than forty years after Lipps introduced the concept in 1909. Within the past decade, however, an upsurge of interest in the measurement of empathy is noticeable in the prodigious research designed to obtain a quantitative index of the empathic process. Rosalind Dymond is responsible for this renewed interest in a personality construct which Speroff (59) characterized as "the concept which has had the dubious distinction of being talked about and written about, but it has never been actually measured."

In 1949 Dymond observed that "although psychologists have touched on the importance of the empathic process, there has been little or no systematic work done on the process itself." Consequently, she attempted to define the field of empathy and to originate a technique to explore and measure the empathic response. Her own definition of empathy; namely, that it is the
"imaginative transposing of oneself into the thinking, feeling, and acting of another" served as the underlying assumption upon which she organized her research endeavor. She assumed that empathic ability is directly proportional to one's ability to take the role of another and make accurate predictions about the way this "other" would respond to a defined task. Thus, empathy could be measured in the laboratory in terms of prediction. She noted that earlier definitions stressed mere identity of feeling and seemed to omit the practical element which to her was a key concept in measuring the empathic response.

In an exploratory study Dymond (20) attempted to apply her assumptions regarding the measurement of the empathic process to some measuring device. She therefore devised a test consisting of four parts, each containing the same six items: self-confidence, superior-inferior, selfish-unselfish, friendly-unfriendly, leader-follower, and sense of humor. A five-point scale was used as the rating measure of each of the six characteristics. The rating procedure for two subjects A and B being tested for their empathy with each other can be demonstrated as follows:

Part I: Subject A rates himself, (A)
Part II: Subject A rates B as he (A) sees him
Part III: Subject A rates B as he thinks B would rate himself
Part IV: Subject A rates himself (A) as he thinks B would rate him

Subject B would go through the same procedure of ratings in reference to Subject A.

A measure of A's empathic ability can be derived by calculating how closely his predictions of B's ratings correspond with B's actual ratings. Similarly, a measure of B's empathy with A can be obtained. This seems like
a logical testing procedure for one who propounds that empathy implies the ability to predict how others will behave in certain defined situations.

Despite its seeming weaknesses which will be discussed later, Dymond's technique of measuring empathy has the distinction of being original. In addition, it is one of two possible paradigms devised to study empathy. Many studies have been patterned after it; and henceforth, the term Dymond-like will be used by this writer in reviewing any study patterned after her method of prediction. The other method of measuring empathy is Kerr's ranking method, which is discussed in Chapter III as part of the historical background of the Empathy Test.

A dichotomy in the research on empathy is discernible mainly on the basis of which testing instrument was used to measure it. Hence, one group of researchers uses the Dymond-like test, either duplicated or modified; the other, Kerr's Empathy Test, the only standardized objective test of empathy.

The main results and conclusions of pertinent studies in each group are brought together in this chapter and their relevance to the present study is noted.

Dymond's 1949 study in which she devised her technique for the measurement of empathy proved to be heuristic in nature. The following findings of that study have given the impetus to extensive experimental endeavors in the field of empathy: (1) A quantitative index of relative empathic ability can be derived from testing instruments; (2) The empathic ability is differentially distributed among people; (3) A significant relationship exists between insight into one's own behavior and empathic ability; and (4) Different personality traits are attributed to high empathizers as opposed to low empathizers.
In another study, Dymond (24) purported to ascertain whether empathic ability increases with age. Each subject was administered two tests which she thought were good measures of empathy. One test was a projective test consisting of pictures which depicted social situations; the other was a social insight test in which the children had to judge which classmates they like and which they dislike. According to Dymond and her associates, both tests required the subjects to assume the role of another. A marked increase with age (from 7 to 11 years) was noticeable in the scores on the projective test. The relationship of insight to popularity measured by sociometry was positive and also increased markedly from the younger to the older age groups.

Dymond (22) also investigated the relationship of empathy to such concepts as sympathy, insight, identification, and projection. Relevant to the study being reported is her finding that good raters seemed to be warm, outgoing, optimistic, emotionally stable people; whereas poor empathizers were rather rigid, introverted people. Although Dymond cautioned that these findings were only preliminary and inconclusive, the writer hypothesized in her study that good adjustment and empathy are positively correlated. Chapter IV discusses the findings regarding this hypothesis.

Another study by Dymond (21) was designed to determine whether clinicians with more empathic ability were in a better position to predict which statements their clients would endorse and which they would reject on the basis of Q-sort. The correlation between counselor's prediction and client's self-sort ranged all the way from .05 to .84. Dymond concluded that counselors having more empathic ability were the ones who made valid predictions.
Arnold and Allport have reservations regarding Dymond's method of measuring the empathic process. For a thorough evaluation of the Dymond studies, the reader is referred to Arnold (3). Allport (2), on the other hand, states that the recent studies purporting to measure empathy are measuring successful judgment, not the particular process of judgment that Lipps had in mind. He concludes that "the theoretical coin has depreciated, probably beyond redemption."

Lindgren and Robinson (45) made an investigation of 45 students who had taken the MMPI as well as the revised form of Dymond's test. Although the relationship between empathy scores and scores received on the various MMPI scales was not thoroughly investigated, it was noted that a group consisting of the "poorest empathizers" received consistently higher (i.e., more maladjustment) scores on the MMPI than did a group consisting of the "best empathizers." These results tend to confirm the general nature of Dymond's (22) findings that persons with "poor" empathy are not as well adjusted according to personality tests as are persons with "good" empathy.

Lindgren and Robinson speculated, however, that a test based on a Dymond-like technique might measure the tendency of individuals to respond to an interpersonal situation in terms of cultural norms rather than empathic promptings. Since both factors seem to operate, they concluded that the present revision of Dymond's test should not be used as a predictive measure of insight or empathy until a form of the test is developed which does not evoke reference to cultural norms.

Bender and Hastorf (8) place the perception of persons at the very heart of social psychology. They administered a scale of 42 items, dealing
with a person's attitudes and feelings toward various situations, to 50 subjects who in Dymond-like fashion attempted to predict the responses of four associates who also responded to the same scale. They concluded that defining empathy by means of a deviational score derived from a Dymond-like test necessarily involved some measure of projection. Therefore, in another study, Hastorf and Bender (34) devised a more sophisticated measure of empathy in which they employed some procedure which took into account projection by the forecaster, as well as similarity between the forecaster and the person whose responses were being predicted. The results showed that the phenomenon of projection was more prevalent than was that of empathy as defined by this study.

Without some correction for projection, according to Hastorf and Bender, attempts to measure empathy do not make psychological sense. Therefore, they emphasized the importance of devising a refined measure of empathic ability, one in which projection is accounted for, to more adequately approximate the psychological aspects of empathy.

With Hastorf and Bender, a change in emphasis became evident in the studies on empathy. Researchers began to regard the empathic process in terms of two behavioral dimensions—interpersonal sensitivity and sensitivity to the generalized other. This shift in emphasis regarding the measurement of empathic responses prompted Kerr and Speroff to devise and subsequently standardize an empathy test based on the assumption that individuals who are superior in empathic ability are persons who are above average in understanding and anticipating the reactions of other people. Chapter III contains a detailed
account of the Empathy Test; the present chapter purports to review some of the studies which have utilized it as a research tool.

Most of the studies on the Empathy Test were of the validation and normative types. During a five-year research program, nine different validation studies have been conducted. In one validation study Bell and Hall (6) indicated that leaders are higher than non-leaders in their ability to identify with the feelings and needs of another. Participants in 18 groups of 5 members each were asked after discussion of a problem, to list their preferences for a leader if the group should meet again. They were also given Dymond's Scale for measuring empathy and Kerr's Empathy Test Form B. The raw data for this experiment consisted of the following information concerning each individual: leadership rankings in each group, Dymond empathy scores based on each five-man group, and Kerr empathy scores which were independent of the groupings.

The correlation between leadership position as measured by peer ratings and empathy as measured by the Kerr Empathy Test was .25. With 81 degrees of freedom this is significantly greater than zero, or better than the 1% level of confidence.

Of great interest to this writer was the near zero correlation between Dymond's and Kerr's tests (r=.02) since both of these tests purport to measure empathic responses. Perhaps the rationale underlying mass empathy for which the Empathy Test was designed is responsible for this low correlation with Dymond's Scales of Empathy which require that the subjects predict some specific person's responses rather than responses to the general population.

The Loyola Language Study requires a subject to predict average persons' responses and in this way seems akin to the Empathy Test.
In another study, a validation attempt by Bell and Stolper (7) purported to investigate the relationship between scores on the Empathy Test and scores derived from the Sensitivity to Other Person Test (STOP). A previous study by Bell and Hall (6) demonstrated that the scores on the Empathy Test are not related to individual empathy as measured by Dymond-like tests of individual empathy. Consequently, the investigators hypothesized that the Empathy Test measures skills related to estimating group opinion since its format requires the subject to predict general attitudes. The STOP test was chosen to test this hypothesis because it requires the subject to estimate the average or group opinion of each member in reference to a series of personality traits. The correlation coefficient between these two tests, however, was not significant, forcing the authors to conclude that their attempt was not successful.

Van Zelst (68) carried out a study relating empathic ability to leadership. Six criteria of leadership ability were obtained from business agents and correlated with their scores on the Empathy Test. The author suggested on the basis of his findings that "the Empathy Test might be profitably employed in the prediction and selection of potential union leaders."

A second validation investigation by Van Zelst (67) resulted in a significant correlation between the Empathy Test scores and two other criteria—job satisfaction and self judgment.

Kerr and Speroff (41) report in their manual for the Empathy Test that empathy is independent of intelligence and the respondent's own feelings. According to them, the test significantly measures a person's feelings for the feelings of others.
In a later study, Speroff (59) set out to determine whether a knowledge of supervisory principles is related to one's ability to assume the role of another individual. Although the population consisted of students, the results appear to be of benefit to industry. They indicate that "one's attitudes, insights, and understanding of the skills necessary for supervisory success, as measured by the test, How Supervise? are significantly related to one's ability to perceive, understand, and anticipate the reactions, feelings, and behaviors of others as measured by the Empathy Test.

Alden (1) conducted a study to determine how a group of people in a clinic-like situation defined their feelings of empathy, and whether their subjective feelings were related to actual similarity to the patient and to ability to predict future behavior of the patients. An attempt was also made to analyze the relationship between self-rated empathy and subjectivity and distortion in the empathizer. The results indicated that self-rated empathy was not significantly related to actual similarity to the patient. Neither was it related to Kerr's Empathy Test significantly. Both Kerr's scores and experience in psychotherapy with the patients, however, were significantly related to accuracy of prediction, (r=.38 and .37).

Rose and Frankel (55) devised a study to determine the relationship of empathic ability as measured by the Empathy Test to the tendency to be chosen in friendship nominations among young teen-agers. In addition, the authors were interested in the chronological course of empathic development. Their study yielded the following results: (1) The relation between empathy scores and sociometric status for these young teen-agers was near zero. This was contradictory to the results of Kerr (41) and Van Zelst (67). The former
reported an \( r \) of .34 between empathy and sociometric choice on college men. The latter found an \( r \) of .59 for the same relationship among skilled building tradesmen. (2) Teen-age males scored higher than teen-age females in average tendencies. This finding was the outcome of a comparison of their data with available data on adults. Their medians of 58 on Form B for boys and 54 for girls compare with the medians of 69 for 117 college men and 57 for 154 college women. The gap in the superiority of teen-age males over teen-age females is narrow in comparison with a similar gap among college students. They, therefore, suggested that this differential gap plus the non-relationship between empathic and sociometric status deserved further investigation.

As can be inferred from the studies reviewed thus far, the results of research in the area of empathy are inconclusive and divergent. Arnold (3) suggests that unless psychologists make

... a clear distinction between empathy as the felt impulse to move as another does, knowing or understanding another as the intuitive and reflective appraisal of this person, and predicting his actions based on an inference from our knowledge of him, very little permanent contribution to knowledge in the field of empathy can be made.

Thus far the review of the literature has been formulated around studies investigating the personality construct—empathy. The following and concluding paragraphs of this chapter are devoted to studies on the Loyola Language Study which measures communality of thought—the other personality concept involved in this study.

The Loyola Language Study made its debut on the psychological scene in 1954. The immediate research following its emergence as a valid diagnostic instrument took the form of normative and validation studies. Later research
endeavors appeared as comparative studies or as correlation studies such as the present one under discussion.

At the time of this writing, thirteen investigations have been completed, either on the master's or doctoral level, using the Loyola Language Study as one of the major research tools. The nature and scope of these studies and their relevance to the present study comprise the subject matter of the following paragraphs.

According to Becker (5) the many research studies on the LLS can be divided into three main groups. Although some studies can be included under more than one classification, the findings are discussed only under the classification that seems most appropriate for his study.

Becker designates the first group as **Normative Studies**: Dinello (17), Guppy (29), Smola (57), and Stanek (60). The findings from these studies form valuable background for the present research because they have pointed out the relationship of such variables as age, sex, education, intelligence, vocational choice, and psychiatric condition to LLS scores. For example, Smola (57) and Guppy (29) both concluded that geographic location has no significant systematic effect upon the test scores. On the basis of their finding, this variable need not be taken into consideration in the present study even though the subjects come from different parts of the country.

At the time of Dinello's study (17), Herr (35) had already substantiated the discriminatory power of the Loyola Language Study between normals and schizophrenics. Details of his study are given in Chapter III as part of the historical background of the LLS. Dinello aimed at investigating whether the LLS scores discriminate subjects within the normal range on the criterion
of occupation. Specifically he investigated whether the scores on the LLS reveal that certain types of personalities are attracted to specific types of jobs. In other words, whether people who seek to enter positions which minimize interaction with others would make significantly different scores than do people who utilize interpersonal relations to a great extent in their jobs. The results indicate that there were trends (low but persistent correlations in the same positive direction) evident in some professional groups. For example, he found that those in managerial and sales positions showed more communality of thought than those in clinical and accounting positions.

In another normative study, Stanek (60) was able to show that age, sex, and education had a significant effect upon the character of responses to the Loyola Language Study. More specifically, he found that age of both sexes bears an inverse relationship to LLS test scores. That is, older persons of both sexes tend to give more unique responses; younger persons more common ones. Education showed a direct relationship to the LLS scores; greater education leads to greater communality of response.

The second group of studies on the Loyola Language Study was categorized by Becker as Predictive Studies: Stewart (61), Del Vecchio (16), and Snider (58).

In one such study, Stewart (61) correlated LLS scores of both men and women with class standing and with the quantitative, linguistic, and total scores for the ACE. No significant relationship was found between any of these scores. Therefore, he concluded that the LLS does not predict college ability or achievement.
Although the present study is not concerned with college achievement, Stewart's findings indicate that the LLS will not be spuriously affected or vitiated by such factors as native intelligence measured by ACE or by diligence in study measured by class standing. This lack of correlation between intelligence scores and LLS scores seems relevant to the present study since the Empathy Test also possesses independence from intelligence and other abilities (Manual).

In this same group of predictive studies, the findings of Del Vecchio (16) in Chicago are most relevant. He found, as also did Stanek and Snider, that the LLS does discriminate between normal and schizophrenic populations both male and female. It seems possible that the LLS may become a single screening test of high validity and reliability in the area of schizophrenia.

The Boston Study (58), which was done prior to the study in Chicago, is described in detail in the next chapter. It also deserves mention in this classification. The findings from this study substantiate Del Vecchio's results regarding the discriminatory power of the Loyola Language Study.

The third group of studies on the Loyola Language Study was classified as Comparative Studies: Braun (9), Logsdon (47), and Trainor (66). The main purpose of two of these studies, (9) and (66), was to compare the responses on the LLS derived by free association (66) or a multiple choice method (9). Trainor's findings support the conclusion that subjects, in general, tend toward significantly greater communality of response through the controlled method which is the unique characteristic of the Loyola Language Study.
Perhaps the research project most responsible for this investigator's desire to work with the Loyola Language Study is Logsdon's (47). The purpose of her study was to determine whether the LLS discriminates between religious women and lay women when matched for age, education, and place of residence. She concluded that the LLS does discriminate elderly lay and religious women. However, younger religious are not discriminated from younger lay women. This finding seems to justify the use of the LLS in a battery of tests for screening candidates to the religious life. Since a nun is first a woman and then a nun, a candidate to the religious life should be a normal young woman differing from other women only in her desire to dedicate her life to God by taking vows and living in a community.

Both Becker (5) and Stewart (62) had the desire to expand the scope of the Loyola Language Study to measures of social perception. Their studies and the writer's are attempts to discover the relationship that might exist between communality of thought and measures of social perception. The writer categorizes these three studies as Correlation Studies. Although all the authors of the studies on the LLS used a correlation technique for one purpose or another to evaluate the results, the distinguishing feature of the present studies is that the LLS scores are correlated with scores on other instruments purporting to measure certain personality variables that involve sociability.

Stewart's study (62) aimed primarily at testing the stability of the Loyola Language Study scores over a rather extended period of time. He, therefore, administered the LLS to the 1958 graduating seniors who had taken the test in 1954 as freshmen. Corollary investigations seemed to suggest
themselves in the light of this primary aim. If such long term stability could be ascribed to the LLS, and if the abiding interests and values of the students could be measured, some kind of relationship should be evident between the scores.

Accordingly, to test the relationship between certain occupational level scores and the LLS scores, Stewart chose Kuder's Preference Record; to test for the relationship between values and LLS scores, he selected the Allport-Vernon-Lindzey Scale of Values.

The correlation coefficient for the four-year test-retest reliability is .68. This finding confirms the reliability reported by Herr (35). However, none of the other correlations is significant even at the .05 level of confidence probably due to the smallness of his samples. Stewart suggests that more research is needed in "the whole area of empathy which is of such supreme importance for studies of human interpersonal relationship."

Herr and other professors carrying Stewart's study further with larger samples cross-validated the study reaching correlations significant at the .05 level of confidence between several interests and scores on the LLS and several values and scores on the LLS. Results were different for men as compared to women, however.

Whereas Stewart sought a relationship to the LLS in the area of occupational interests and social values, Becker turned to the particular area of group social perception designated as perception to the generalized other in order to seek possible relationships that would enhance our understanding of the Loyola Language Study.
By administering the Loyola Language Study simultaneously with a classical measure of social sensitivity and observing the similarities and differences between the two measures, he aimed at showing empirically that the LLS was or was not a measure of social sensitivity.

Although Becker's study involves six hypotheses, only one has direct relevance to the present study. The writer, therefore, limits herself to a consideration of this relevant hypothesis. According to Becker (5)

A Loyola Language Study score is actually an accuracy score measuring the accuracy with which an individual perceives the general public. It does this by measuring his ability to estimate what words are most frequently associated to those in the stimulus list. As such an accuracy score it should be positively correlated with other accuracy scores of this nature which have been designated to measure this same ability to perceive the general public.

After considering all available personality inventories, Becker chose items from The Guilford-Zimmerman Temperament Survey (GZTS) to obtain scores with which to correlate the Loyola Language Study scores. In order to introduce the element of prediction into the GZTS so that the basic set required of the subjects for responding would be similar for both instruments, he requested the subjects to answer the questions on the GZTS and then to predict the percentage of the group that would answer the question in the same manner. The resulting score from this prediction according to Becker, would be designated as the Social Perception Accuracy Score, although much doubt exists as to the validity of his method.

The correlation between the Loyola Language Study scores and the Social Perception Accuracy scores is -.02 showing essentially no more than chance relationship. Despite this lack of correlation and probably because
Becker doubts the validity of his method of "percentage of group" estimate, he insists that "the Loyola Language Study is a measure of social perception." And in order to encourage further research in this area he says of the LIS: "Its assets and liabilities are best understood in the theoretical framework of social perception measures. Improvements and meaningful refinements are likely to result from research in that area."

The writer shares Becker's conviction regarding the theoretical framework of the Loyola Language Study. Consequently, the present study has been formulated to ascertain the relationship between the Loyola Language Study and another measure from the area of social perception--Kerr's Empathy Test.
CHAPTER III

METHODOLOGY–DATA COLLECTION

Subjects

The subjects for this study are 118 young religious women ranging in age from 17 years to 27 years. Mean age is 20–5 and Standard Deviation is 2 years. Table 1 categorizes them into their respective status groups in religious life. It is pertinent to a better understanding of the rating procedure, which is discussed at the end of the chapter, if the reader be cognizant of this classification.

TABLE 1

TABULATION OF SUBJECTS ACCORDING TO MEMBERSHIP GROUPS,
MEAN GROUP AGE, AND TIME IN RELIGION

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean Age</th>
<th>Time in Religion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior Postulants</td>
<td>20</td>
<td>17-7</td>
<td>1 month</td>
</tr>
<tr>
<td>Senior Postulants</td>
<td>16</td>
<td>18-8</td>
<td>1 year</td>
</tr>
<tr>
<td>Novices</td>
<td>17</td>
<td>20-3</td>
<td>2 years</td>
</tr>
<tr>
<td>Neo-Professed</td>
<td>17</td>
<td>21-4</td>
<td>3 years</td>
</tr>
<tr>
<td>Junior Professed</td>
<td>18</td>
<td>21-5</td>
<td>4 years</td>
</tr>
<tr>
<td>Junior Professed</td>
<td>14</td>
<td>22-8</td>
<td>5 years</td>
</tr>
<tr>
<td>Junior Professed</td>
<td>16</td>
<td>24-1</td>
<td>6 years</td>
</tr>
</tbody>
</table>
That nuns function as subjects for this research endeavor is not fortuitous. The implications of empathic ability and communality of thought are directly applicable to nuns whose active apostolate and daily community living demand continuous empathic involvement and constant dynamic interaction with people. The investigator was interested, therefore, in determining the empirical relationship between the criterion of adjustment to community living and the scores on the Loyola Language Study and the Empathy Test. Consequently, it seemed feasible to select a non-random sample of young religious because of the nature of the hypotheses under investigation.

THE RESEARCH INSTRUMENTS

The experimental design of this study required three research instruments—the Loyola Language Study, the Empathy Test, and an Adjustment Rating Scale. The underlying assumptions of the Loyola Language Study and the Empathy Test have been briefly considered in Chapter I.

The purpose of the present chapter is threefold: (1) to discuss the historical backgrounds, formats, and scoring systems of these testing instruments, (2) to delineate the testing procedure involved in this study, and (3) to describe the statistical treatment of the data.

The Loyola Language Study

The origin of the Loyola Language Study and its subsequent screening efficiency are directly related to the Boston Study conducted in 1953 by Louis B. Snider and Olof Johnson at the Boston State Hospital. Consequently, any authentic description of the historical background of the Loyola Language
Study necessarily commences with a discussion of the procedure and the results of that psychological research endeavor.

The most significant development attributable to the Boston Study is the emergence of a new association technique empirically established as a valid diagnostic psychiatric procedure. This new association technique requests the subject to respond with the one word which he believes most other people would give to a stimulus word. To do this, he must deliberate. Thus, the new technique is a modification of the traditional set which requires the subject to respond to a stimulus word with the first word that comes to his mind.

Although it is generally easy for a normal individual to recognize if his associations are in conformity with those of the general population, it is also plausible for a neurotic to perform as a clinically normal person on an association test. A possible cause of this spurious result may be found in the psychic defenses which a neurotic is prone to utilize. Thus, Guppy (29) states that the merely neurotic realizing that his unique ideas and strong impulses are different, acts not so much as sheer impulse dictates, but through the exercise of control, performs in patterns of behavior adaptable to what is commonly done.

The psychotic, however, usually characterized by strange associations, eccentric thinking, inability to perceive what normal behavior is, and a lack of control over the externalization of his psychic life, unwittingly betrays himself in his highly individualized responses. Thus, Snider and his associates were convinced that the demands of their new association technique
could not be fulfilled by psychotics. Moreover, they predicted that any instrument employing the technique would prove highly discriminatory, especially in disorganized states of greater severity, such as schizophrenia. Their prediction was shared by other researchers on word association tests and sustained by the Kent-Kosanoff studies from which emerged the conclusive finding that "there is a gradual and not an abrupt transition from the normal to the pathological state."

In order to confirm their conviction with empirical evidence, Johnson and Snider undertook the task of using the new technique to establish norms against which the performance of the patients could be compared. They administered their 80-word test to a sample of 400 males and 400 females from the Boston area. The subjects were stratified according to age, sex, and education because at that time it was not yet known whether these variables affected the scores on an association test. These 800 tests, representing a stratified random sample, constituted the first set of norms for subjects in the Boston metropolitan area. Later Stanek (60) established norms on an equal number of subjects in the Chicago metropolitan area, and Guppy (29) did the same for the Seattle metropolitan area.

With norms established, Snider and Johnson subsequently attempted to validate the theory underlying their new technique by administering the 80-word test to 78 female psychotic patients at the Boston Hospital. The results showed that 70% of all the hospitalized schizophrenic females fell below the 10th percentile of the normal sample, and with one exception, all the patients fell below the 15th percentile. Their prediction that controlled association
as employed in the Loyola Language Study discriminates significantly between normals and hospitalized schizophrenics was thus corroborated.

Del Vecchio (16) administered the LLS to 56 male and 53 female schizophrenics in the Chicago area. The results also strongly confirmed the LLS as a valid discriminatory tool between schizophrenics and normals.

In 1954, Loyola University, Chicago, copyrighted the original testing material used in the Boston Study under the name Loyola Language Study. Since then, twelve years of validation and refinement have enhanced its position among association tests.

The present format of the Loyola Language Study is a 5½ by 8½ four-page booklet easy to handle and non-threatening in appearance. It can be administered individually, in groups, or be self-administered. Testing time ranges from 20 to 35 minutes. Its contents consist of the 80 words chosen from the Kent-Rosanoff list of 100 words which Snider and Johnson used for normative and discriminatory purposes in Boston.

A glance at the directions on the first page of the LLS testing booklet reveals that two key phrases (one word and greatest number of people) are set apart from the rest of the wording in bold print and repeated five times for emphasis. (See Appendix I.) The emphasis is not superfluous, however, since these two key phrases are the verbal expression of the new association technique originated by Snider and Johnson. Consequently, they embody the principle of control which distinguishes the Loyola Language Study from older traditional controlled association tests.

Other controlled association tests require the subject, after sufficient reflection, to give a subordinate or supracoordinate to the concept
contained in a stimulus word. It cannot be overemphasized that the Loyola Language Study is not a controlled association test in this same sense. It is rather a semi-controlled association test demanding not a categorical response but a deliberative choice of the one word that the subject believes the majority of the population would give as a response to the stimulus word. It rather resembles the Minnesota test for "popular responses" recently used by Jenkins with the same 100 original words of the Kent-Rosanoff list.

There are three scoring methods which have been formulated to derive scores on the Loyola Language Study: (1) the standard score method, (2) the double-root frequency method, and (3) the median score method. According to Herr (35) insignificant differences result from the different methods of scoring association tests. He reports correlations as high as .96 between standard scores and double-root scores, and .93 between standard scores and median scores. In addition, any of the three scoring systems yields significant differences between the group means of psychotics and normals beyond the .001 level of confidence.

Since no statistical analysis of LLS scores can be made without direct reference to the scoring system used to derive them, a brief consideration of the salient features of each method is presented as basic to a meaningful discussion of the Loyola Language Study as a psychological testing instrument.

The original scoring method was the standard score method. One of the first steps involved in its formulation was the establishment of standard score scales. Herr (35) devised such scales on data derived from the Boston Study and the Chicago Study by Stanek.
The standard score method, costly in time and effort, consisted essentially in assembling a frequency count for each response to every one of the 80 stimulus words. It was assumed that there was a continuum from high frequency to low running parallel to and corresponding to the strength of association bonds between a stimulus and response thereto. The frequencies were then converted into appropriate percentages from which the square roots were extracted in order to reduce the length of the distribution.

This procedure, of converting frequencies to percentages and then extracting the square roots, resulted in a psychological continuum based on the responses of the normative group. Accordingly, the higher frequencies denote stronger stimulus-response connections, and vice-versa.

After the mean and standard deviation for each stimulus-response distribution had been calculated for all the 800 Boston subjects as well as for the 800 Chicago subjects, standard scores (which now comprise the standard score scales) were computed with an arbitrary mean set at 20 and a standard deviation at 10. The sum of the standard scores for all 80 stimulus-response pairs was computed to yield a single score for each subject. This single score, designated as the subject's communality score, was considered indicative of the person's over-all success on repeated trials in finding common responses.

Although the standard score method had its advantages, a simpler method, called the double-root frequency method, was later devised by Doyle and Stanek. This method is based on the same notion of reducing the range of frequencies and gives to any given response a value absolutely determined by
its raw frequency, but does not reduce the frequency means and deviations to any kind of standard scores.

After the booklets of the Chicago sample of 800 normals and 109 patients were rescored according to these absolute root-frequency scales, it was found that in some cases the variances were proportional to the means. This precluded the use of certain statistical procedures with the data.

In a final effort to find the best system of scoring the LLS, Rimoldi* was consulted. He collaborated to devise the median score system. It involved the assignment of a score of "1" or "0" to each response, depending on whether it met the "pass" or "fail" category.

This method was utilized on the same 800 Chicago normals and 109 Chicago patients used in establishing the other scoring methods. The first step involved the location of the response of the highest frequency to each stimulus word. As many other response frequencies (in descending order) were accumulated to the highest frequency until 50% of the population was included. All responses falling within or above this median frequency merited a score of "1". Any other response not included in the 50% cumulative frequency scored "0".

The median score method, however, did not seem practicable after it was found that many abnormals scored a zero in over half of the items on the shortened form of the Loyola Language Study. Therefore, median scores were not adequate in producing fine discriminations in levels of abnormality.

*Director of the Psychometric Laboratory, a division of the Psychology Department of Loyola University.
In a compilation of the results of all previous investigators, Herr matched the following subjects for age and education allowing one year of difference for either criterion: 53 Chicago female patients with 53 Chicago female normals; 56 Chicago male patients with 56 Chicago male normals; 50 Boston male patients with 50 Boston normal males; and 50 Boston female patients with 50 Boston normals. The results showed that the mean difference between normals and patients in each area was significant beyond the .001 level of confidence.

A closer look at the above results suggested the hypothesis that the significant differences emanate from definite stimulus-response combinations. Consequently, the items in the booklets of these same matched subjects were sifted by Herr and his associates, and those 25 items which contributed the most to the difference between normals and psychotics were isolated. The scores for each of these 25 items for normals were significantly different at the .01 level of confidence from the scores of abnormal for the same items.

However, it is evident from Table 2 that only 11 of the top distinguishing items are common to both Chicago men and women, and therefore, scoring norms are necessary for males and females separately.

After the booklets were rescored using only these 25 highly discriminating words, and after new means and standard deviations were computed, it was concluded that two advantages accrue from the use of the shortened form of the Loyola Language Study: (1) The screening efficiency of the test is enhanced; and (2) The scoring time is considerably reduced.
### TABLE 2
THE LOYOLA LANGUAGE STUDY. DISCRIMINATING WORDS FOR CHICAGO MALES (M) AND FEMALES (F)*

<table>
<thead>
<tr>
<th>soldier (F)</th>
<th>high (F)</th>
<th>working</th>
<th>sickness (M)</th>
<th>dream (F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hungry</td>
<td>white</td>
<td>comfort</td>
<td>mountain (M, F)</td>
<td>trouble (F)</td>
</tr>
<tr>
<td>butterfly</td>
<td>command</td>
<td>soft (M)</td>
<td>stove (F)</td>
<td>religion (F)</td>
</tr>
<tr>
<td>long</td>
<td>sour</td>
<td>short</td>
<td>girl</td>
<td>street</td>
</tr>
<tr>
<td>head (M, F)</td>
<td>king (M)</td>
<td>beautiful</td>
<td>salt</td>
<td>health</td>
</tr>
<tr>
<td>anger</td>
<td>deep</td>
<td>cold (M)</td>
<td>man</td>
<td>ocean</td>
</tr>
<tr>
<td>afraid</td>
<td>sleep (F)</td>
<td>whiskey</td>
<td>cheese</td>
<td>bed (F)</td>
</tr>
<tr>
<td>fruit</td>
<td>black (M)</td>
<td>yellow</td>
<td>baby</td>
<td>child</td>
</tr>
<tr>
<td>dark (M)</td>
<td>hammer (F)</td>
<td>window (M, F)</td>
<td>moon (F)</td>
<td>tobacco (M, F)</td>
</tr>
<tr>
<td>red</td>
<td>table (M, F)</td>
<td>scissors (M)</td>
<td>spider</td>
<td>woman</td>
</tr>
<tr>
<td>loud (F)</td>
<td>thirsty (M, F)</td>
<td>foot (M, F)</td>
<td>bread (M, F)</td>
<td>cabbage</td>
</tr>
<tr>
<td>bath</td>
<td>quiet</td>
<td>doctor (M, F)</td>
<td>whistle (M, F)</td>
<td>citizen (F)</td>
</tr>
<tr>
<td>eating</td>
<td>hard</td>
<td>wish (M)</td>
<td>carpet</td>
<td>earth</td>
</tr>
<tr>
<td>joy (M)</td>
<td>blue</td>
<td>house (F)</td>
<td>needle (M)</td>
<td>lion</td>
</tr>
<tr>
<td>rough (F)</td>
<td>sweet</td>
<td>justice</td>
<td>hand (M)</td>
<td>butter (M)</td>
</tr>
<tr>
<td>heavy</td>
<td>stomach (M, F)</td>
<td>river</td>
<td>thief (M)</td>
<td>music</td>
</tr>
</tbody>
</table>


Two essential characteristics of the Loyola Language Study remain to be discussed: (1) its reliability and (2) its validity.

Reliability indices on the LLS have been reported either in terms of coefficients of internal consistency or coefficients of stability. For example, Herr (35) correlated the odd-numbered items with the even-numbered items on samples of 400 men and 400 women. The resulting split-half reliabilities corrected by the Spearman-Brown formula for attenuation ranged from .88 to .94.

Using the same split-half method, Smola (57) computed coefficients of internal consistency ranging from .72 to .89.
Trainor (66) reports a test-retest coefficient of stability of .49 between scores obtained on the same normal persons who took the retest 4 to 8 weeks later. However, when only the 25 top screening words were scored for these same individuals, the coefficient increased to .55.

Undoubtedly, the most important question that needs to be raised regarding any psychological test concerns its validity. Three methods were used by Herr (35) to estimate the screening efficiency of the Loyola Language Study. One coefficient of validity for the total scores on the shortened test was computed by a method devised by Eysenck. In practice, it involved subtracting the proportion of normals incorrectly identified, from the proportion of patients correctly identified by the LLS. Indices of .64 for Chicago men, .70 for Chicago women, .70 for Boston men, and .71 for Boston women resulted.

The other method involved the conversion of Phi coefficients, computed on each of the 25 items, into Product-Moment coefficients. The assumption of continuity between normals and patients on communality of thought served as the basis. The resulting median coefficients were .45 for Chicago men, .66 for Chicago women, .65 for Boston men, and .68 for Boston women.

Another attempt at assessing the validity of the Loyola Language Study was made by Herr after the death of Father Snider. It required a correlation between the ratings of psychologists and psychiatrists on 50 Boston female patients and their scores on the 25 top screening words. A Product-Moment coefficient of .48 resulted which is significant beyond the .01 level of confidence.

These reported validity indices indicate that the Loyola Language Study is characterized by concurrent validity, which, according to the
Technical Recommendations for Psychological Tests and Diagnostic Techniques

is inherent in any psychiatric screening instrument that discriminates between identifiable groups.

The Empathy Test

Psychologists have conceived and constructed tests for the measurement of virtually all the abilities, aptitudes, and interests known to man. An objective test, however, for the measurement of the rather unique personality construct called empathy was not forthcoming until 1947 when Williard B. Kerr devised the Empathy Test—the first objective group-type test of empathic ability. (See Appendix II.)

In 1951, after undergoing a five-year research program, the Empathy Test emerged revised, standardized, and validated. In 1961 another revision of the test resulted from the collaborative efforts of Kerr and Speroff. The present study has utilized the 1961 revised adaptation of the Empathy Test, Form B.

According to Kerr, a valid measure of empathic ability makes provision for the criterion of prediction. He, therefore, limited his choice of test items to those stimulus situations having universal appeal and ample opportunity for the respondent to predict typical reactions of defined normative persons. Consequently, any item which was the obscure psychological possession of a little known ethnic, regional, or socio-economic group was automatically excluded on the basis of not meeting the criteria of universality and predictability.
After having sifted the various possible item sources, Kerr formulated the Empathy Test around the categories of music, reading, and interpersonal relations because he felt that these three areas not only possessed the greatest universal appeal, but also had empirically established normative data.

Part I is constructed in the area which philosophers sometimes call the common language of mankind--MUSIC. The types of music were chosen on the advice of a sales executive for RCA Victor Phonograph Records. Included in the listing are the following types: classical, semi-classical, waltzes, Hawaiian, blues, polkas, fast dance, western, square dance, hill-billy, humor novelty, "Hit Parade" type, sacred, and spirituals. The normative data supporting this section resulted from a national survey program.

Part II is formulated around the reading and buying preferences of all types of magazines. Consequently, it potentially represents the general interests and aspirations of a cross section of the citizenry and calls for a realistic empathic identification with special classes and interest groups.

Included in the list of fifteen current magazines are popular and well-known ones like Life, Newsweek, McCalls, Harpers, American Home, Field and Stream, Fortune, and New Yorker; as well as less popularly known publications such as Stamps, Argosy, Ebony, Pageant, Cosmopolitan, and Southern Farm and Home. Supporting normative data are based on the actual sales of such publications.

Part III is organized with special reference to the area of annoying interpersonal relations. Loud chewing of gum, affected manner, and habitual arguing are but a few of the listed annoying behavior patterns. Normative reports are based on the extensive findings of Hulsey Cason (41).
The nature of the response is the same for each part. Consequently, the subject ranks from most to least the popularity of the music types, the most widely read magazines, or the annoying experiences.

The "set", however, that the subject must affect while responding to the test items is different from each part. Hence, in Part I the testee identifies with a hypothetical office worker of the United States, and is asked to predict his work music desires by giving a rank of "1" to the music type he thinks the typical office worker prefers while working, a rank of "2" to the next preferred, and down the line to the least preferred. (See Appendix II.)

In Part II the subject shifts his identity to the typical American reader of magazines and is requested to rank the magazines in order from most to least paid circulation.

In Part III the subject assumes the role of a typical person over 40 and is required to rank the annoying experiences by giving a rank of "1" to the experience he feels a person over 40 would consider most annoying, "2" to the next most annoying, etc. It is obvious that the format requires the subject to predict general attitudes. The reader will recall that the format of the Loyola Language Study also requires response to the general population. Hence the similarity between the two instruments.

Unlike the scoring methods of the Loyola Language Study, the scoring system of the Empathy Test is relatively simple involving no complicated statistical procedures. It is based on a Ranking Key which was formulated from normative data derived from a national survey program for Part I, the actual sales of publications for Part II, and the extensive findings of
Hulsey Cason for Part III. Specifically, the computation of the Empathy Score consists in summing the differences (regardless of sign) between the rank assigned by the testee and the ranking that is given by the Key for each item. The sum is subtracted from an arbitrary value of 200, and this remainder comprises the Raw Score.

The Empathy Test Manual contains substantiative evidence on reliability. Test-retest reliability coefficients for Form A ranging from .67 (108 college men) to .80 (60 eastern college students) are reported. Van Zelst (67) obtained a split-half reliability coefficient of .89 (corrected via Spearman-Brown formula) for Form A on 124 male skilled workers. Equivalent form reliability of .83 determined by intercorrelations between Forms A and B on 74 college men with approximately one year intervening between administrations is recorded.

The fundamental prerequisite of any research instrument is its validity. Nine validation studies have been conducted using the Empathy Test as the main research tool. For the results of these studies the reader is referred to Chapter II where these studies have been reviewed as part of the related literature.

The Adjustment Rating Scale

It was hypothesized in this study that religious rated as well-adjusted by their associates are more likely to have desirable scores on the Loyola Language Study and the Empathy Test, than those religious rated as poorly adjusted or moderately adjusted. Since a numerical index of adjustment was needed in order to make the comparison, the writer devised an Adjustment
Rating Scale based on the criterion of adjustment to community living. A three-point rating system was used as the scale.

The procedure involved was rather simple. Thus, the rater was requested to assess a subject's overt adjustment to community living by checking one of the following descriptive units that might apply to the subject being rated: (1) poorly-adjusted; (2) moderately-adjusted; (3) well-adjusted.

The instructions to the rater are as follows:

You have been living in close association with Sister ________ for ____ years. As far as her adjustment to community living, what kind of nun do you feel sister is? Check one.

____ Poorly-adjusted
____ Moderately-adjusted
____ Well-adjusted

In order to give the raters a frame of reference, "well-adjusted" was defined as "fits in well into the various situations of life including social situations as well as one's academic environment; with old people as well as with people of her own age." With this operational meaning as a gauge, the other two judgments could be made according to the amount of deviation from the rating, "well-adjusted."

Because the criterion of adjustment is not a numerically measured variable on the Adjustment Rating Scale described above, a proper system of weighting had to be devised which would convert the frequencies of the different rating units to one numerical index amenable to statistical treatment. Consequently, weights of 300, 200, and 100 were adopted. The weights were assigned to the three rating units in descending order beginning with the criterion—"well-adjusted."
To depict the manner in which the numerical adjustment index was computed by this weighting system, the following example is offered.

Subject A belongs to a group having sixteen members. She receives ratings of 5 well-adjusted, 10 moderately adjusted, and 1 poorly adjusted. Her total weighted score amounts to \((5 \times 300) + (10 \times 200) + (1 \times 100)\) or 3600. Dividing this by 16 yields an average score of 224 which is considered the numerical index of her adjustment to community living as evaluated by this study.

It was assumed that the ratings were made honestly and seriously. If this assumption is correct, the group average rating represents a valid description of how an individual appears to those with whom she interacts dynamically. Coming from a source external to the subject's self evaluation, this group rating can be looked upon as the most objective criterion of how well a subject gets along with her social environment.

According to Brownfain (10) so intimate is the relationship between self and the social environment, that hardly any maladjustment exists in the personality that is not somehow reflected in maladjustment in interpersonal relations. One would, therefore, look to the group's evaluation of an individual as a crucial criterion of her adjustment. Any bias emanating from unique experiences or the "halo" effect is minimized by the fact that the average rating is a combination of 13 or more judgments.

The foregoing factual information relevant to the testing instruments has been offered as a necessary frame of reference for a broader understanding of the testing procedure which is described below.
A preliminary preparation for the actual testing included the division of the 118 subjects into workable testing groups. Group membership in religious life seemed to be the best criterion to follow. (See Table 1.) Consequently, the Junior and Senior Postulants (N=36) constituted one group; the Novices and Neo-Professed (N=34) another group; and all the Junior Sisters (N=48) constituted the third.

Since three research instruments are utilized in this study, provision had to be made for more than one testing session. Accordingly, it was deemed feasible to administer the Loyola Language Study and the Empathy Test in one session and the Adjustment Rating Scale in another. To insure uniformity of procedure, all the tests were administered by the investigator.

The test administration was preceded by a brief orientation-motivation period. In order to establish good rapport between the investigator and the subjects and to encourage wholehearted cooperation, the subjects were assured that their scores would have no bearing on their present status in the community. They were also guaranteed complete anonymity by a system of coding. Such a protective device seemed necessary to encourage sincerity and a willingness on the part of the subjects to give realistic and unbiased ratings of their group members on the Adjustment Rating Scale.

Personal data on the subjects were obtained in the following manner. Three-by-five file cards bearing a code number were distributed, and the subjects were asked to supply the information requested. The items included name, age, birth, years in aspirancy, years in religion, and educational level. This information would be used only in case the analysis of results
was contingent upon specific information of this nature. Each subject was instructed to identify all her tests with the code number found on her file card.

With preliminaries completed, the Loyola Language Study booklets were distributed. The examiner read the LLS instructions orally emphasizing the key phrases one word and most people would think of. Then the subjects were asked to reread the directions carefully and to indicate whether or not they understood them. Although no time limit was set, all the subjects finished within 35 minutes; some as early as 25 minutes.

Upon completion of the Loyola Language Study, the subjects were given a short intermission after which the Empathy Test was distributed. The procedure was duplicated. The examiner read the instructions orally, but this time emphasis was placed on the particular identification that the subject had to assume before responding to the items. It will be recalled that the subject shifts identity for each of the three parts. The subjects were advised to reread the instructions before responding to the items. Although no time limit was set, most of the subjects completed the test within 25 minutes.

The Adjustment Rating Scale was administered the following day. The directions called for the subject to rate the members of her group according to the criterion of adjustment to community living. Because it is essential to know a person before passing judgment on his behavior, the subjects were asked to rate only members of their groups. (See Appendix III for sample of the Rating Scale.) Consequently, the Postulants rated Postulants, the Novices
rated Novices, etc. The reader is referred to Table 1 for a review of the categorized groups.

Each subject was given as many copies of the Rating Scale as there were members in her group. It was felt that if each subject were rated by all her group members and the mean rating calculated, an adequate measure of the group consensus would be obtained. The mean rating score was computed according to the weighting system devised for this study to establish a numerical index of adjustment.

STATISTICAL TREATMENT OF THE DATA

The research data assembled from the testing instruments just described were subjected to the following statistical techniques: The 118 Loyola Language Study booklets were scored according to standard z-scores on the 25 most discriminating words. These z-scores constitute the standard score scales which were derived from the responses of the normative sample of 400 Chicago women. The responses were arranged on a descending scale according to the per cent of agreement of any respondent in choosing a particular response. Lower scores indicate better adjustment (as is true of the MMPI) and higher scores indicate poorer adjustment.

The Empathy Test Scores were derived from the 1962 revision of the Key. Raw scores were evaluated in terms of the normative table established for Liberal Arts Women reported in the Manual. (See Appendix IV.) Chosen from twelve other normative tables, this one seemed most apropos, since the subjects in the present study were pursuing a Liberal Arts Program.
Separate means and standard deviations were computed for the distributions of the Loyola Language Study scores, the Empathy Test scores, and Mean Rating scores. It was necessary to compute the means and sigmas for these distributions because a desirable score area is usually delimited by two sigmas above and below the mean of a distribution. A desirable score area was needed to test Hypothesis II.

The present investigation appears to be the first experimental analysis of the relationship that exists between the Loyola Language Study and a standardized measure of empathic ability such as the Empathy Test. It is also a first attempt to investigate the relationship of Loyola Language Study scores and Empathy Test scores to Adjustment Indices. The Pearson Product-Moment Correlation technique was chosen as the most appropriate statistical method to explore these relationships since the data fulfill the requirements for this statistical procedure. Consequently, the following three Pearson Product-Moment correlations were computed: (1) on the Loyola Language Study scores with Empathy Test scores; (2) on Loyola Language Study scores with Adjustment Indices; and (3) on Empathy Test scores with Adjustment Indices.
CHAPTER IV

ANALYSIS OF THE RESULTS

The present study is primarily an experimental analysis of the relationship between the scores on the Loyola Language Study and Kerr's Empathy Test. In the light of this analysis, two hypotheses are proposed: (1) There is a positive correlation between desirable Loyola Language Study scores and desirable Empathy Test scores, and (2) Religious rated as well-adjusted by their associates are more likely to have desirable scores on the Loyola Language Study and the Empathy Test than those religious rated as poorly adjusted or moderately adjusted.

The 118 LIS scores in the present research range from 399 to 773 with a Mean of 527.5 and a Sigma of 77. These scores were derived from the standard score scale for Chicago adult women devised by Herr. His normative scale yielded a Mean of 500 and a Sigma of 80. The rationale underlying his scoring procedure implies that a score of 500 is normal; 580 is one Sigma worse; 420 one Sigma better, etc. It is obvious, that according to this scoring method, lower scores are desirable and denote less schizoid tendency.

A range of desirable scores had to be delimited in order to test Hypothesis II. In the previous chapter, it was established that the limits for a desirable score area would be designated by 2 Sigmas above and 2 Sigmas below the Mean. On the present population of N=118, this range includes LIS
scores from 373.5 to 681.5. Of the 118 subjects, 112 obtained scores within this range, while 6 received scores falling 2 Sigmas above the Mean and are, therefore, designated as undesirable scores. The range of the six undesirable scores is 688 to 773. The histogram in the Appendix presents a graphic representation of the LLS frequency distribution for this population.

The Empathy Test scores of the 118 subjects range from 28 to 112 with a Mean of 73.4 and a Sigma of 16.17. According to the norms for Liberal Arts Women provided in the Manual for the test, this Mean falls at the 80th percentile. (See Appendix IV.) In contrast to the LLS scores, high Empathy Test scores denote more empathic ability and are, therefore, desirable scores.

Table 3 shows graphically the comparison of the subjects in this study with the norms for Liberal Arts Women.

**TABLE 3**

<table>
<thead>
<tr>
<th>%tile</th>
<th>Norm Score</th>
<th>N</th>
<th>%tile</th>
<th>Norm Score</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>107</td>
<td>1</td>
<td>45</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td>95</td>
<td>83</td>
<td>32</td>
<td>40</td>
<td>53</td>
<td>6</td>
</tr>
<tr>
<td>90</td>
<td>81</td>
<td>10</td>
<td>35</td>
<td>51</td>
<td>2</td>
</tr>
<tr>
<td>85</td>
<td>75</td>
<td>14</td>
<td>30</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>80</td>
<td>73</td>
<td>6</td>
<td>25</td>
<td>46</td>
<td>4</td>
</tr>
<tr>
<td>75</td>
<td>71</td>
<td>8</td>
<td>20</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td>70</td>
<td>69</td>
<td>7</td>
<td>15</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>65</td>
<td>67</td>
<td>8</td>
<td>10</td>
<td>38</td>
<td>2</td>
</tr>
<tr>
<td>60</td>
<td>64</td>
<td>8</td>
<td>5</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td>50</td>
<td>60</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The norms were abstracted from the Manual for the Empathy Test.*
A glance at Table 3 reveals that 71 subjects fall at the 75th percentile or better, 28 subjects at the 50th percentile or better, and 13 subjects at the 25th percentile or better.

From Table 4, it can be seen that when the above totals are converted to percentages, 60% of the subjects fall above the 75th percentile. On the basis of the percentages shown in Table 4, the writer concludes that a favorable comparison with the normative sample is evidenced by the subjects of this study.

**TABLE 4**

**NUMBER AND PERCENT OF SUBJECTS ACHIEVING SCORES AT Q3 (OR ABOVE), Q2 (OR ABOVE), Q1 (OR ABOVE), AND BELOW Q1 BASED ON NORMS IN TABLE 6**

<table>
<thead>
<tr>
<th></th>
<th>Q3 (or above)</th>
<th>Q2 (or above)</th>
<th>Q1 (or above)</th>
<th>Below Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Subjects</td>
<td>71</td>
<td>28</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Percent of Group</td>
<td>60%</td>
<td>24%</td>
<td>10%</td>
<td>6%</td>
</tr>
</tbody>
</table>

The same procedure for establishing a desirable range of scores was followed for the Empathy Test scores as had been for the Loyola Language Study. Consequently, the Mean, plus and minus 2 Sigmas, comprises the desirable score area. Scores of 41.6 to 105.74 fall within this area. Of the 118 subjects, 113 received desirable scores ranging from 28 to 40. The histogram in the Appendix provides a graphic picture of the Empathy Test frequency distribution for the subjects in this study.
Table 5 presents the number of desirable and undesirable scores on the Loyola Language Study and on the Empathy Test. The data reveal that the results of both tests are quite similar.

**TABLE 5**

RANGE OF DESIRABLE AND UNDESIRABLE SCORES ON THE LOYOLA LANGUAGE STUDY AND THE EMPATHY TEST FOR 118 SUBJECTS

<table>
<thead>
<tr>
<th>Description of Limits</th>
<th>Loyola Language Study</th>
<th>N</th>
<th>The Empathy Test</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desirable Scores</td>
<td>375.5-681.5</td>
<td>112</td>
<td>41.6-105.74</td>
<td>113</td>
</tr>
<tr>
<td>Undesirable Scores</td>
<td>688-773</td>
<td>6</td>
<td>28-40</td>
<td>5</td>
</tr>
</tbody>
</table>

The Rating Averages (derived by the method described in Chapter III) range from 121 to 300 with a Mean of 242 and a Sigma of 38. The histogram in the Appendix discloses that the Rating Averages approximate a normal distribution.

The Rating Averages are divided into three areas according to the three descriptive rating units (well-adjusted, moderately-adjusted, poorly-adjusted). Thus, scores of 120 to 199 are designated as "poorly-adjusted" scores; 200 to 274 as "moderately-adjusted" scores; and 275 to 300 as "well-adjusted" scores. Table 6 shows the number of subjects falling within each category of scores. The implication of the rating results seems to indicate that, on the whole, these young religious women impress each other favorably.
### Table 6

RANGE OF SCORES FOR EACH DESCRIPTIVE CATEGORY ON THE ADJUSTMENT RATING SCALE

<table>
<thead>
<tr>
<th>Category</th>
<th>Score Limits</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-adjusted</td>
<td>275-300</td>
<td>23</td>
</tr>
<tr>
<td>Moderately-adjusted</td>
<td>200-274</td>
<td>81</td>
</tr>
<tr>
<td>Poorly-adjusted</td>
<td>120-199</td>
<td>14</td>
</tr>
</tbody>
</table>

The ranges of scores, the Means, and the Sigmas of all three instruments used in this study are accumulated in Table 7.

### Table 7

MEANS, SIGMAS, AND RANGES OF SCORES ON THE LOYOLA LANGUAGE STUDY, THE EMPATHY TEST, AND THE ADJUSTMENT RATING SCALE

<table>
<thead>
<tr>
<th>Description</th>
<th>The Loyola Language Study</th>
<th>The Empathy Test</th>
<th>The Rating Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of scores</td>
<td>399-773</td>
<td>28-112</td>
<td>120-300</td>
</tr>
<tr>
<td>Mean</td>
<td>527.5</td>
<td>73.4</td>
<td>242</td>
</tr>
<tr>
<td>Sigma</td>
<td>77</td>
<td>16.17</td>
<td>38</td>
</tr>
</tbody>
</table>

The preceding paragraphs report only the resulting scores of the three instruments used. A discussion of the implications of these data is the subject matter of the following paragraphs.

Table 8 provides the basis for testing Hypothesis II. Thus, of the 118 subjects, 23 received ratings designated as "well-adjusted" scores. All
23 subjects also obtained desirable Loyola Language Study scores and desirable Empathy Test scores. This finding seems to substantiate Hypothesis II which states that subjects rated as well-adjusted by their peers are more likely to have desirable LLS and ET scores than subjects rated as moderately adjusted or poorly adjusted.

### TABLE 8

**A COMPARISON OF THE FREQUENCY COUNT FOR EACH RATING CATEGORY WITH FREQUENCY COUNT OF DESIRABLE AND UNDESIRABLE SCORES ON THE LOYOLA LANGUAGE STUDY AND THE EMPATHY TEST**

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Loyola Language Study Scores</th>
<th>The Empathy Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Desirable</td>
<td>Undesirable</td>
</tr>
<tr>
<td>Well adjusted</td>
<td>23</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Moderately adjusted</td>
<td>81</td>
<td>76</td>
<td>5</td>
</tr>
<tr>
<td>Poorly adjusted</td>
<td>14</td>
<td>13</td>
<td>1</td>
</tr>
</tbody>
</table>

However, when the other two categories are taken into consideration, no consistent pattern of a relationship between the scores is discernible. For example, of the 81 subjects receiving ratings amounting to "moderately adjusted" scores, 10 received undesirable scores, either on the LLS or on the ET. In no instance were the 5 subjects who received undesirable LLS scores, the same subjects who received undesirable Empathy Test scores. Of the 14 subjects rated as poorly-adjusted, 13 of them received desirable LLS scores, and all 14 achieved desirable ET scores. Hence, no pattern of a relationship is evident.
The data were finally analyzed in terms of Pearson Product-Moment correlations, and the implications of the results are discussed below.

A Pearson Product-Moment correlation was computed on the Loyola Language Study with the Empathy Test. A resulting $r$ of -.22 proved significant at the .02 level of confidence. The negative correlation, however, spuriously indicates that a subject high on empathy is low on communality of thought and vice versa. This apparent negative correlation is, in reality, positive. We know from the assumptions underlying standard scores that a high score on the Loyola Language Study is undesirable. The opposite must be said of the Empathy Test on which a high score is desirable. With this reservation kept in mind, the writer concludes that the data support Hypothesis I which states that a positive correlation exists between the Loyola Language Study scores and the Empathy Test scores. An empathic person, therefore, is also one who shares common associations with the general population.

In order to verify Hypothesis II through a statistical procedure, Pearson Product-Moment correlations were computed on the Loyola Language Study scores with the Rating Averages, as well as on the Empathy Test scores with Rating Averages. A correlation coefficient of -.30 significant at the .01 level of confidence was obtained between the LLS and Rating Averages. However, an insignificant $r$ of -.03 resulted between the ET scores and Ratings Averages.

For a graphic presentation of these correlation coefficients, the reader is referred to Table 9.
TABLE 9

PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN SCORES
ON THE LOYOLA LANGUAGE STUDY AND THE EMPATHY TEST;
ON THE LOYOLA LANGUAGE STUDY AND THE RATING SCALE;
ON THE EMPATHY TEST AND THE RATING SCALE

<table>
<thead>
<tr>
<th></th>
<th>LLS</th>
<th>ET</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loyola Language Study</td>
<td>-.22**</td>
<td></td>
<td>-.30*</td>
</tr>
<tr>
<td>The Empathy Test</td>
<td>-.22</td>
<td></td>
<td>-.03</td>
</tr>
<tr>
<td>The Rating Scale</td>
<td>-.30</td>
<td></td>
<td>-.03</td>
</tr>
</tbody>
</table>

*Significant at the .01 level of confidence
**Significant at the .02 level of confidence

The significant negative correlation between the LLS scores and the Rating Averages must be interpreted with the same reservation as the negative but significant correlation between ET and LLS scores. Consequently, this writer concludes that the data seem to support Hypothesis II, that there is a general and significant tendency for subjects with better adjustment to have better LLS scores. The question as to why the hypothesis holds for the Loyola Language Study but not for the Empathy Test cannot be answered satisfactorily at this point.
CHAPTER V

CONCLUSION

The present research was designed primarily to determine empirically the relationship existing between the scores on the Loyola Language Study and the scores on Kerr's Empathy Test.

A secondary purpose aimed at ascertaining whether any patterned relationship could be detected among the scores on the Loyola Language Study, the Empathy Test, and the Adjustment Rating Scale.

Two hypotheses were proposed: (1) There is a positive correlation between desirable Loyola Language Study scores and desirable Empathy Test scores. (2) Religious rated as well-adjusted by their associates are more likely to have desirable scores on the Loyola Language Study and the Empathy Test than those religious rated as poorly-adjusted or moderately-adjusted.

The subjects of this study were 118 young religious women ranging in age from 17 to 27 years with a Mean Age of 20.5 and a Sigma of 2 years.

The testing procedure involved the administration of the Loyola Language Study and the Empathy Test in one testing session, and the Adjustment Rating Scale in another.

The historical backgrounds of the Loyola Language Study and Kerr's Empathy Test were traced from matrix to current status as psychological testing
instruments. The salient features of each test, including underlying assumptions, scoring methods, validity, and reliability, were delineated. Special emphasis was given to the derivation of the scores on the Loyola Language Study, the Empathy Test, and the Rating Scale which was devised for this study.

A perusal of the literature revealed that no study had been conducted utilizing the Loyola Language Study and the Empathy Test in the same research endeavor.

From the viewpoint of the personality constructs measured by each test, it seemed practicable to utilize both instruments in the same research design. The Empathy Test purports to measure mass empathy or a subject's ability to predict the average person's responses. The Loyola Language Study, on the other hand, assesses communality of thought or a subject's ability to predict the responses most people would make to a given stimulus word. Mass empathy and communality of thought are related personality constructs since both responses appear to be reactions to the generalized other. It has been proven by this study that a relationship does exist between them.

In Chapter II, it was pointed out that the research studies on empathy were dichotomized on the basis of the procedure used to measure the empathic process. Consequently, Dymond's predictive technique and Kerr's ranking method were evaluated as two paradigms devised thus far to measure empathy. The results of one study using both methods revealed that no empirical relationship existed between Kerr's Empathy Test and Dymond's Empathy Scales. This failure to find a relationship between the two methods seems to indicate
that a difference exists between mass empathy as measured by the Empathy Test and individual empathy as measured by the Dymond Scales.

Valuable background information was gleaned from the review of the research on the Loyola Language Study. For example, the findings of these studies have pointed out the relationship of such variables as age, sex, education, intelligence, vocational choice, and psychiatric condition to scores on the Loyola Language Study.

The statistical techniques applied to the data included the computation of the means, standard deviations, and Pearson Product-Moment correlations. The means and standard deviations were used to delimit the desirable score ranges for the Loyola Language Study, and the Empathy Test scores. A comparison of the Loyola Language Study scores and the Empathy Test scores with the Adjustment Rating Averages revealed that no consistent pattern of a relationship existed among them. However, when these same scores were analyzed on the basis of a Pearson Product-Moment correlation, a different picture presented itself.

Thus, a resulting r of -.30 between the Loyola Language Study and the Rating Averages proved to be significant at the .01 level of confidence. The negative correlation has to be interpreted as positive in this manner: The more desirable is the Loyola Language Study score, the better is the rated adjustment.

The r of -.22 resulting between the Loyola Language Study scores and the Empathy Test scores proved to be significant at the .02 level of confidence. The negative correlation, however, is misleading in view of the
meaning that is attached to the numerical value of the scores on the Loyola Language Study and the Empathy Test. We pointed out that a high score on the Empathy Test and a low score on the Loyola Language Study are desirable. Consequently, the resulting r indicates that a person who has the ability to match his associations successfully against the common associations shared by the general population is also a person who can predict normative responses of people by identifying with a normative person.

With the above interpretation of the negative correlation coefficients kept in mind, the writer concludes that both of the hypotheses in the study have been substantiated by the results.

There is some basis in the literature on empathy for two assumptions that seem to be relevant to the present study. (1) There is a significant relationship between empathic ability and insight into one's own behavior. (2) The concept of self formed by a well-adjusted person harmonizes quite accurately with objective observations others make of him.

The investigator suggests that the verification of these assumptions be a problem for future research. In addition, a corollary investigation designed to ascertain the relationship of insightful perception of self to the scores on the Loyola Language Study seems feasible.


LOYOLA LANGUAGE STUDY

Instructions

When people see or hear a word, they often think of another word. If you say the word stem, most people would think of flower. Some, but not the greatest number, might think of pipe, grass, stop, and so forth.

This study wants to find out what word you think the greatest number of people would be most likely to think of when they see or hear each of the words on the next two pages.

Please write next to each of the words the one word which you think the greatest number of people would be most likely to think of when they see or hear the word in the list. Take as much time as you need to think about the word which seems to you to “go along” with each printed word. Then choose the one word which you think the greatest number of people would be most likely to think of when they see or hear the given word. Write the one word which you choose beside the printed word. Do not skip any word.

Remember, you are not asked to write down just any word that comes to your mind. You should write down the one word which you think the greatest number of people would be most likely to think of.

Important: please fill out the information blank on page 4.

Copyright 1954, by LOYOLA UNIVERSITY, CHICAGO
DIRECTIONS

How well do you know the likes and dislikes of average people? In the test items below, try to place yourself in the position of the hypothetical person described. Answer the questions in such a way as to agree with the actual facts, were they available.

1. Below are fourteen common types of music. Rank them in order of their popularity among the office workers of the United States. Give a rank of "1" to the most popular, "2" to the second most popular, etc., and "14" to the least popular.

   RANK MUSIC
   1 polkas
   2 classical
   3 blues
   4 waltzes
   5 humor-novelty
   6 fast dance
   7 Hawaiian
   8 square dances
   9 western
   10 sacred
   11 spirituals
   12 "Hit Parade" type
   13 hill-billy
   14 semi-classical

2. Below are the names of fifteen magazines. Rank them in order from most to least paid circulation.

   RANK MAGAZINE
   1 Stamps
   2 Harper's
   3 McCall's
   4 Nation
   5 New Yorker
   6 Southern Farm and Home
   7 Newsweek
   8 Ebony
   9 Argosy
   10 American Home
   11 Fortune
   12 Pageant
   13 Life
   14 Field & Stream
   15 Cosmopolitan

3. Below are ten commonly annoying experiences to persons aged over 40. Imagine yourself a typical person of this age level and rank from most to least annoying the following.

   RANK ANNOYING EXPERIENCE
   1 A person constantly looking glum
   2 Hearing a person chewing gum
   3 A person habitually arguing
   4 A person with an affected manner
   5 A person not noticing what I say
   6 Body odor
   7 A person with a gushing manner
   8 A person being too inquisitive
   9 Coaxing me to do something I don't want to do
   10 Being told to do something just as I am about to do it

Copyright 1951 - W. A. Kerr
(Reproduction by any means strictly prohibited.)

Published by Psychometric Affiliates

Box 1625
CHICAGO 1, ILL.
APPENDIX III

THE ADJUSTMENT RATING SCALE

ADJUSTMENT RATING SCALE

You have been living in close association with Sister _______ for _____ years.

As far as her adjustment to community living, what kind of a nun would you say Sister is?

Check one.

_____ Poorly-adjusted     _____ Moderately-adjusted     _____ Well-adjusted
APPENDIX IV

NORMS FOR THE EMPATHY TEST
## Norms

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Union Business Members</th>
<th>College Students</th>
<th>Industrial Supervisors</th>
<th>College Men</th>
<th>Retail Clerks</th>
<th>Chicago Firemen</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>163</td>
<td>102</td>
<td>126</td>
<td>117</td>
<td>115</td>
<td>122</td>
</tr>
<tr>
<td>95</td>
<td>157</td>
<td>93</td>
<td>111</td>
<td>105</td>
<td>108</td>
<td>115</td>
</tr>
<tr>
<td>90</td>
<td>151</td>
<td>92</td>
<td>107</td>
<td>98</td>
<td>104</td>
<td>112</td>
</tr>
<tr>
<td>85</td>
<td>147</td>
<td>88</td>
<td>99</td>
<td>96</td>
<td>101</td>
<td>107</td>
</tr>
<tr>
<td>80</td>
<td>144</td>
<td>83</td>
<td>90</td>
<td>91</td>
<td>95</td>
<td>103</td>
</tr>
<tr>
<td>75</td>
<td>139</td>
<td>81</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>95</td>
</tr>
<tr>
<td>70</td>
<td>136</td>
<td>78</td>
<td>87</td>
<td>85</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>65</td>
<td>125</td>
<td>76</td>
<td>86</td>
<td>81</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>60</td>
<td>121</td>
<td>74</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>55</td>
<td>118</td>
<td>73</td>
<td>84</td>
<td>80</td>
<td>71</td>
<td>78</td>
</tr>
<tr>
<td>50</td>
<td>112</td>
<td>71</td>
<td>82</td>
<td>65</td>
<td>73</td>
<td>65</td>
</tr>
<tr>
<td>45</td>
<td>107</td>
<td>70</td>
<td>80</td>
<td>72</td>
<td>61</td>
<td>68</td>
</tr>
<tr>
<td>40</td>
<td>103</td>
<td>69</td>
<td>78</td>
<td>69</td>
<td>56</td>
<td>63</td>
</tr>
<tr>
<td>35</td>
<td>100</td>
<td>66</td>
<td>75</td>
<td>65</td>
<td>51</td>
<td>58</td>
</tr>
<tr>
<td>30</td>
<td>98</td>
<td>65</td>
<td>70</td>
<td>63</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>25</td>
<td>97</td>
<td>62</td>
<td>63</td>
<td>60</td>
<td>41</td>
<td>48</td>
</tr>
<tr>
<td>20</td>
<td>87</td>
<td>59</td>
<td>60</td>
<td>58</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>15</td>
<td>84</td>
<td>54</td>
<td>54</td>
<td>51</td>
<td>31</td>
<td>38</td>
</tr>
<tr>
<td>10</td>
<td>76</td>
<td>46</td>
<td>45</td>
<td>45</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>74</td>
<td>35</td>
<td>39</td>
<td>20</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>72</td>
<td>12</td>
<td>34</td>
<td>15</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

N = 64 124 52 110 432 65 64 55

References: grateful acknowledgements are due to 1. Frank J. Holmes; 2. Frank J. Smith; 3. Raymond H. Van Zeist.

## Form A

### Engineering Men Combined Scores Form A PLUS B Administered in Succession

<table>
<thead>
<tr>
<th>A+B</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>241</td>
<td></td>
</tr>
<tr>
<td>194</td>
<td></td>
</tr>
<tr>
<td>190</td>
<td></td>
</tr>
<tr>
<td>177</td>
<td></td>
</tr>
<tr>
<td>174</td>
<td></td>
</tr>
</tbody>
</table>

## Form B

### Hospital Auto Indust Volunte Sales Man Super visors

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Junior High School Seniors</th>
<th>College Students</th>
<th>Industrial Supervisors</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>98</td>
<td>88</td>
<td>114</td>
</tr>
<tr>
<td>95</td>
<td>93</td>
<td>76</td>
<td>102</td>
</tr>
<tr>
<td>90</td>
<td>60</td>
<td>74</td>
<td>98</td>
</tr>
<tr>
<td>85</td>
<td>78</td>
<td>78</td>
<td>94</td>
</tr>
<tr>
<td>80</td>
<td>75</td>
<td>75</td>
<td>89</td>
</tr>
<tr>
<td>75</td>
<td>71</td>
<td>73</td>
<td>86</td>
</tr>
<tr>
<td>70</td>
<td>70</td>
<td>70</td>
<td>84</td>
</tr>
<tr>
<td>65</td>
<td>66</td>
<td>69</td>
<td>80</td>
</tr>
<tr>
<td>60</td>
<td>61</td>
<td>67</td>
<td>78</td>
</tr>
<tr>
<td>55</td>
<td>60</td>
<td>63</td>
<td>76</td>
</tr>
<tr>
<td>50</td>
<td>58</td>
<td>61</td>
<td>73</td>
</tr>
<tr>
<td>45</td>
<td>57</td>
<td>61</td>
<td>73</td>
</tr>
<tr>
<td>40</td>
<td>56</td>
<td>57</td>
<td>70</td>
</tr>
<tr>
<td>35</td>
<td>55</td>
<td>54</td>
<td>68</td>
</tr>
<tr>
<td>30</td>
<td>45</td>
<td>52</td>
<td>66</td>
</tr>
<tr>
<td>25</td>
<td>44</td>
<td>50</td>
<td>63</td>
</tr>
<tr>
<td>20</td>
<td>42</td>
<td>49</td>
<td>63</td>
</tr>
<tr>
<td>15</td>
<td>39</td>
<td>47</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>44</td>
<td>56</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>18</td>
<td>40</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>-2</td>
<td>24</td>
</tr>
</tbody>
</table>

N = 35 47 22 43 53 414 62 101 98 64 32 93
APPENDIX V

--- Mean = 527.5
Sigma = 77

Fig. 1. Distribution of scores of 118 subjects on the Loyola Language Study.
APPENDIX VI

---

Fig. 2. Distribution of scores of 118 subjects on Kerr's Empathy Test.
Fig. 3. Distribution of scores of 118 subjects on the Adjustment Rating Scale.
The thesis submitted by Sister Mary Lucinia Szpak, C.S.S.F. has been read and approved by the director of the thesis. Furthermore, the final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the thesis is now given final approval with reference to content, form, and mechanical accuracy.

The thesis is therefore accepted in partial fulfillment of the requirements for the Degree of Master of Arts.

Feb 2, 1967
Date

Signature of Adviser