



1965

## Clinical Judgment as a Function of Manifest Anxiety and Social Conditions/

Joseph F. Pribyl  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_theses](https://ecommons.luc.edu/luc_theses)



Part of the [Psychology Commons](#)

---

### Recommended Citation

Pribyl, Joseph F., "Clinical Judgment as a Function of Manifest Anxiety and Social Conditions/" (1965).  
*Master's Theses*. 2028.

[https://ecommons.luc.edu/luc\\_theses/2028](https://ecommons.luc.edu/luc_theses/2028)

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).  
Copyright © 1965 Joseph F. Pribyl

**Clinical Judgment as a Function of  
Manifest Anxiety and Social Conditions**

by

**Joseph F. Pribyl**

**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**

**January**

**1965**

## Vita

Joseph F. Pribyl was born in Chicago, Illinois, on October 29, 1940.

He graduated from St. Ignatius High School, Chicago, in June, 1958. He received the degree of Bachelor of Science in the Natural Sciences from Loyola University, Chicago, in June, 1962.

He began his graduate studies in the Department of Psychology of Loyola University in September, 1962. From then until June, 1964 he was an assistant in the perception laboratory at Loyola. He has been a Psychologist Trainee at the Reception and Diagnostic Center of the Illinois Youth Commission in Joliet, Illinois, since June, 1964. He is presently taking his clerkship at the above installation.

### Acknowledgements

The author wishes to express his gratitude to Dr. Ronald E. Walker for his help and encouragement during the preparation of this thesis. He is also indebted to Dr. William A. Hunt of Northwestern University for his encouragement and monetary assistance for this study as part of a larger project being conducted through contract Nonr 1228 (18) with the Office of Naval Research.

## Table of Contents

	Page
Introduction . . . . .	1
Method . . . . .	17
Results . . . . .	21
Discussion . . . . .	30
Summary . . . . .	35
References . . . . .	37
Appendix A (Data Sheet) . . . . .	40
Appendix B (Questionnaire). . . . .	42
Appendix C (Mean Interjudge Reliabilities and the Results of Duncan's New Multiple Range Test). . . . .	44

List of Tables

Table	Page
1. Significance of Differences Between Mean Interjudge Reliabilities for 16 High Anxious (HA) and 16 Low Anxious (LA) <u>Ss</u> in Group and Individual Conditions . . . . .	23
2. Significance of Differences Between Mean Interjudge Reliabilities for 16 High Anxious (HA) <u>Ss</u> in the Group (G) Condition and 16 HA <u>Ss</u> in the Individual (I) Condition and for 16 Low Anxious (LA) <u>Ss</u> in the G Condition and 16 LA <u>Ss</u> in the I Condition . . . . .	24
3. Analysis of Variance of the $z'$ Values for the Mean Interjudge Reliabilities for the Four Groups of Subjects . . . . .	26

## List of Figures

Figure	Page
1. Mean interjudge reliability (in $z'$ values) on a set-by-set basis. . . . .	27
2. Mean interjudge reliabilities (in $z'$ values) of the four experimental groups combined for each set of stimuli and the mean standard deviations of the clinicians' ratings of each set of stimuli. . . . .	29

Clinical Judgment as a Function of  
Manifest Anxiety and Social Conditions

Joseph F. Pribyl

Loyola University, Chicago

Historical Background. Because experienced clinicians so frequently are required to make diagnoses with little information and little time, the feeling has developed that clinical judgment was an intuitive process. This feeling brought with it the implication that clinical judgment had a mystical quality that made it unapproachable by ordinary scientific means. Going against this tradition, Sarbin, Taft, and Bailey (1960) produced a clinical judgment model based on the processes of syllogistic reasoning. There are several stages in the process of clinical judgment according to their analysis; namely, the development of a postulate system in the person doing the judging, the eduction of premises, the establishment of cues and their use to instantiate the object, and the drawing of conclusions from the instantiation in terms of the predicate of the major premise (Sarbin et al., 1960, p. 20). Taking another approach, Hoffman (1960) adopted a mathematical model based on information theory. Both of these approaches present difficulties that hinder fruitful research. Sarbin et al., have given a good rational analysis of clinical judgment, but they have not given much in the way of testable hypotheses (Hunt & Jones, 1962). While hypotheses are forthcoming from Hoffman's mathematical model, they are not presently testable because available analyses of clinical judgment have not identified with sufficient precision the cues or inputs that are



pertinent to an information theory model.

An earlier attempt by Meehl (1954) seems to offer a somewhat more hopeful approach. He considers the processes involved in clinical judgment as analogous to those involved in actuarial prediction. Meehl suggests that making statistical predictions on the basis of actuarial tables is a more exact way of doing much of what a clinician does in making intuitive decisions. The clinician has a finite number of facts that he puts together in different combinations of importance to make predictions. He also possesses a series of "rules of thumb" that he uses in making decisions. The operations that a clinician goes through in making a decision based on a set of facts can be done by a clerical worker, a calculator, and actuarial tables. The actuarial method is likely to be more accurate in predicting because the method assigns the weights that are optimal for best predictions to the different facts. While Meehl favors the use of the superior actuarial method and its high predictive value, he realizes that even if vast actuarial tables and techniques were available they could not replace the clinician in the creative act of making a hypothesis. Hunt and Jones (1962) state that the actuarial method is theoretically the best method of clinical diagnosis; but they realize that at the present time the actuarial method is not the answer to the problems of clinical diagnosis. Preventing the fulfillment of the actuarial method's promise of accuracy is the fact that the actuarial approach is useful only in areas where refined tests are available. The actuarial approach is also hampered by public opinion, which objects to

the use of machines in making judgments about men.

Development of the Psychophysical Model. Because of these difficulties, Hunt (1959) feels that the presently used clinical methods should be improved through research. One of the difficulties encountered in doing research on clinical judgment is that any given judgment is based on a unique set of facts that cannot be reproduced. Underwood (1957) points out that one of the requirements for scientific investigation is a reliable phenomenon. The kind of clinical judgment that occurs in daily clinical practice would seem to lack this prerequisite of reliability and thus not be amenable to scientific study. Hunt (1959) suggests that if the clinician making repeated clinical judgments were made the focus of clinical research, clinical judgment could be made the subject of scientific investigation.

Every clinical judgment has its unique aspects, but each also shares certain commonalities with other judgments, particularly those made by the same judge. These commonalities can be the subject of rigorous scientific investigation as the determinants of individual judgments by a single clinician. By the same token, the variables that influence agreement among several judges can be studied by comparing judgmental performance in identical, or at least similar, situations. This concept of interjudge agreement forms the basis for much of Hunt's work.

Hunt (1959) has suggested that the situation in which several clinicians are asked to make repeated judgments on the same clinical material is analagous to the paradigm of classical psychophysics. In his

work Hunt uses the method of single stimuli by having clinical material rated along some scale. It is hoped that clinical judgment can be shown to be one of several phenomena embodied within the general category of judgment. If this is true, then much of the literature pertaining to psychophysical judgment can be brought to bear on the problem of clinical judgment.

In the context of the psychophysical model, Hunt and Arnhoff (1956) have demonstrated that clinical judgment is reliable as measured by inter-judge agreement. Other workers (Campbell, Hunt, & Lewis, 1957; Campbell, Lewis, and Hunt, 1958) have shown that the context effects well known to classical psychophysics (Beebe-Center, 1929; Helson, 1947; Hunt, 1941; Hunt & Volkman, 1937; and Johnson, 1955) are also found in clinical judgment.

In classical psychophysics, variables that are logically related to the field of learning have been shown to affect judgment. Helson (1947, 1948) has shown that the Ss' previous acquaintance with similar stimuli changes the Ss' adaptation level (a phenomenon in which perception of previous stimuli will influence perception of subsequent stimuli). It would seem that experienced clinicians should be better able to make clinical judgments than naive judges since they have had experience with a wider range of stimuli. Several investigators (Grigg, 1958, Hunt, Jones, & Hunt, 1957; Jones, 1957; Cline, 1955) have confirmed the above.

Learning Theory and Judgment. In efforts to relate clinical judgment to other areas of psychology, Hunt and his co-workers have begun to investigate the relationship between clinical judgment and learning theory. In doing this Hunt and Jones (1962) hope that clinical judgment will become

more firmly anchored in experimental psychology. Gibson (1953) reviewed many studies showing that absolute judgments made with the method of single stimuli improve when there is only practice but no correction or knowledge of results. Ammons (1955) reviewed experiments dealing with different types of judgments and perceptual-motor performances and concluded that learning is faster and reaches a higher level with knowledge of results and that the more specific the knowledge the more the rapid the improvement. In consideration of the above evidence Blumberg (1961) predicted that (a) practice in making clinical judgments with no knowledge of results would lead to more rapid, reliable, and valid judgments; (b) even more rapid, reliable, and accurate clinical judgments would result if the judge were given specific knowledge of the correct judgmental responses; (c) clinical judgments of an intermediate degree of rapidity, reliability, and validity would result if only general feedback were given to the judges; and (d) there would be transfer of training (greater rapidity, reliability, and validity of judgments) when new stimuli were judged. Having Ss rate the vocabulary responses from hospitalized schizophrenics on a 7-point scale of exhibited disorganization, Blumberg found that the three conditions made no difference in the rapidity of the judgments, and that hypothesis (a) above was not supported in that the reliability and the validity of the clinical judgments did not improve when the judges received only practice and no feedback, but the reliability and validity of the clinical judgments did improve when the judges received the general and specific feedback as predicted in hypotheses (b) and (c) above. The hypothesis that there

would be transfer of training in all three conditions was not supported in that transfer was found only in the condition in which the judges received specific feedback.

One of the more vigorous areas of research in the field of learning has been the concept of drive as measured by anxiety scales (Sarason, 1960). Taylor (1951, 1953, 1956) developed the first anxiety scale to receive widespread attention. Taylor's Manifest Anxiety Scale (hereafter referred to as MAS) was originally designated as an operational measure of Hull's drive in an eyelid conditioning experiment (Taylor, 1951). Taylor developed the Hullian based hypothesis that different sources of drive summate in Ss to produce a total effective drive state (D) that sets the strength of the conditioned eyelid response. Taylor assumed that different levels of psychiatrically defined "manifest anxiety" would be indicative of different levels of generalized drive. She obtained 65 true-false items which 80 percent of a group of clinical staff members chose as being indicative of manifest anxiety as it was operationally defined. The 65 selected items were part of a group of 200 MMPI items that the clinicians judged. The original MAS items were mixed in with 135 MMPI items not related to anxiety.

Taylor's original scale was later (1953) cut to 50 items that showed the highest correlation with the total score, and these 50 items were mixed with the L, K, and F scales of the MMPI and MMPI items scored on Wesley's rigidity scale. The final scale numbered 225 items and has been called the Biographical Inventory.

Taylor (1951) found that high anxious Ss (Ss scoring high on the MAS)

were consistently superior to low anxious Ss (Ss scoring low on the MAS) in the amount of eyelid conditioning (hereafter high anxious Ss will be referred to as HA and low anxious Ss as LA). The results were statistically significant. An attempt, through two sets of differential instructions after 20 eyelid conditioning trials, to induce experimentally differing levels of stress in the Ss failed to produce any statistically significant differences. Taylor interpreted the differential eyelid conditioning obtained for the two groups of Ss selected on the basis of their MAS scores as meaning that the drive level of the HA Ss was higher than that of the LA Ss and hence that the growth curves of the excitatory potentials for the two groups of Ss were different. Taylor also suggested that on the basis of Hull's (1943) postulate that the growth of excitatory potential was dependent upon both habit strength (H) and drive (D), the difference in the growth curves of excitatory potential in the two groups (inferred from differences in the conditioning curves) might be due to changes in both D and H. In such a case, the HA Ss would react more strongly to the unconditioned stimulus implying that the same physical stimulus had a different psychological value for the HA Ss and LA Ss. Taking into consideration Hull's (1943) postulate that reward partially determines H, the termination of the unconditioned stimulus should produce a greater reduction of D in the HA Ss, and, therefore, increase H.

While higher drive level (inferred from higher MAS scores) should lead to better performance in a situation where there is only one habit evoked, the predictions for tasks in which there are several available habits having

differing levels of availability are more complex. Taylor (1956) suggested that in a complex task two other Hullian (1943) concepts must be used. They are oscillatory inhibition (O) and threshold (L). The following characteristics are attributed to O: (a) O varies from moment to moment such that the distribution of O for a group of individuals on the same response at any moment would be approximately normal; (b) O plays an inhibitory role, subtracting from excitatory potential and thus giving rise to momentary excitatory potential. For a given response to occur, the momentary excitatory potential must be higher than the threshold value (L) for that response. It is assumed that the value of L is the same for like habit tendencies evoked in a particular situation. In a task where several response tendencies are available in competition, the one that will take place is the one with highest momentary excitatory potential. Keeping in mind the postulate that excitatory potential is dependent upon habit strength, other things being equal, the response tendency with the greatest H and therefore the greatest excitatory potential has the greatest probability of taking place. Adding the conception of D as affecting excitatory potential, when the desired response is weaker (lower H) than one or more competing response tendency, the SS with higher D will perform less well than SS with lower D. One further possibility exists in that responses having very weak habit strengths may gain enough excitatory potential to be above threshold, thus reducing the probability of the correct response in the high D SS. In the case where the correct response is maximally available, heightened drive would make performance superior for high drive SS.

While Blumberg (1961) had established that learning (improvement in reliability and validity) did take place when Ss had general and specific feedback, there were indications that different kinds of learning took place even with no feedback. For example, rating the same stimuli (schizophrenics' vocabulary test responses) over six trials reduced the latencies over trials of the Ss' judgments even with no feedback. This finding was replicated in another study (Hunt and Blumberg, 1961). If nothing else, the Ss were learning their own judgmental responses better.

Anxiety and Clinical Judgment. The question arose as to just when a subject, in making repeated judgments, is judging evaluatively and when he is simply repeating previous responses. The assumed parallel to the S's experimental judging is that of a clinical situation in which a practicing clinician gets faster and faster in making clinical evaluations. When does the clinician stop making clinical, judgmental evaluations and simply start repeating previously learned responses to relevant stimuli? An attempt was made to answer the above question for the Ss making experimental judgments by applying Taylor's drive theory to the task of repeated clinical judgments. Ss who score high on the MAS should initially perform less well than Ss who score low on the MAS. The difference in performance of high and low scorers on the MAS should shrink with repeated judgments and they should perform equally well. According to drive theory (Taylor, 1956), those Ss scoring higher on the MAS would have a greater response probability for competing responses, thus making incorrect responses more likely. However, once the high MAS scorers establish the correct response, they should perform with



shorter latency than low MAS scorers. The disorganizational cues upon which the rating of the schizophrenic's test responses is based would provide the competing response tendencies in the above formulation. The point at which the performance curves for the high MAS scorers and low MAS scorers would cross, as predicted by drive theory, would be the point at which evaluative judgment stopped and the elicitation of learned verbal responses began. Hunt and Blumberg (1961) had high MAS scorers and low MAS scorers rate 21 schizophrenics' vocabulary test responses on a 7-point scale of disorganization in different orders over six trials. The measures of learning were latency, the number of shifts in judgment, reliability or interjudge agreement, and validity as represented by the agreement of the judge with the standardized values of the stimuli. All four measures indicated that learning took place. Only the reliability and validity measures, however, differentiated the high MAS scorers from the low MAS scorers, with the low MAS scorers being superior to the high MAS scorers on trial one and the differences diminishing by the sixth trial. The performance curves of the two groups of Ss did not cross, thus placing this particular application of Taylor's drive theory in doubt.

As a check on the results of the Hunt and Blumberg (1961) study, Hunt and Walker (1963) reanalyzed the data with a trial-by-trial analysis and obtained a significant difference between the HA Ss and LA Ss only on the first trial. Hunt and Walker also exactly replicated the Hunt-Blumberg study with a new set of subjects. The results paralleled the reanalysis of the Hunt-Blumberg study except for what was probably a chance difference

between the HA Ss and LA Ss on trial two.

To check the possibility that there were not enough trials to permit the crossing of the performance curves of the HA and LA Ss, Hunt and Walker (1963) did a second study utilizing 100 different standardized schizophrenic test responses presented in 10 sets of 10 stimuli equated in range of standardized stimulus values. While the Hunt and Blumberg study (1961) and Hunt and Walker's (1963) replication of it demonstrated that HA and LA Ss were differentiated in performance on repeated judgments of the same stimuli, the use of 100 different stimuli permitted the researchers to find out if the performance differences of HA and LA Ss would also be present if only a general frame of reference was learned. Hunt and Walker's second experiment (1963) showed that only on the first set of 10 stimuli did LA Ss perform better than the HA Ss with the two groups of Ss being equally reliable on the remaining nine sets of stimuli. Because the results of three different studies did not support Taylor's (1956) drive theory in its prediction of the crossing of the performance curves of the HA and LA Ss, Hunt and Walker (1963) suggested that what Child describes as "irrelevant responses made to anxiety" (1954, P. 154) were greater for Ss who scored high on the MAS than for Ss who scored low on the MAS, and that the HA Ss eliminated the task irrelevant responses quickly, allowing their performance to come up to that of the LA Ss.

Social Situation and Clinical Judgment. Many of the clinical judgment studies done by Hunt and/or his co-workers dealt with subjects and experimenters in a one-to-one relationship. Walker, Hunt, and Schwartz (in press)

have integrated Child's (1954) interpretation that HA Ss have more task irrelevant responses with a discussion of the relation of stress and task irrelevant responses presented by Spence (1963). They applied their integration to the comparison of Ss making clinical judgments in a co-acting, non-interacting group and Ss making clinical judgments in an individual (or one-to-one) situation.

Walker et al. (in press) had noted that there was apparent lessening of tension for Ss in experiments on judgment if the Ss judged in a group rather than individually. Spence (1963) suggests that the intensity of task irrelevant responses is related to the amount of stress in an experimental situation. On the basis of the above, Walker et al. predicted that the HA Ss would have a lowered or equal intensity of task irrelevant responses relative to LA Ss when both judged in a group situation. Such a difference between Ss working in a group and individually was assumed to be due to the existence of comparatively less stress in a group clinical judgment experiment as compared to a clinical judgment experiment in which there is a one-to-one relationship between the E and the S.

In three independent experiments that utilized a group testing situation, the above conclusions were supported (Walker et al., in press). In two experiments there were no significant differences between HA and LA Ss over many clinical judgments. In the remaining experiment the HA Ss were superior to the LA Ss in early judgments but not in later ones.

Allport (1920, 1924) made clinical observations that appear to be in contrast to those that Walker et al. (in press) reported. In describing

the individual differences among the Ss who worked in groups and relating the individual differences to their experimental performances on non-judgmental tasks, Allport remarked that some "nervous" Ss were not helped by working in a group but were hindered. Allport did not have an objective measure of "nervousness". If what Allport called nervousness were assumed to be a drive characteristic possessed by HA Ss, it could be suggested that HA Ss would have poorer judgmental ability in a group situation. This paradox may be explained when one considers that Allport's observations about nervous individuals were made on Ss performing non-judgmental tasks and experimental group situations that many Ss described as competitive.

In the experiments done by Walker et al. the task was a judgmental situation which would be unlikely to produce competition among the Ss. Thus, the variable of competition or no competition among Ss might account for the disparate results.

In an unpublished study, Pribyl (1963) had two random groups of naive Ss rate 50 schizophrenic vocabulary responses on a 7-point scale of disorganization. One group was tested individually and the other was tested in a co-acting group. The 50 stimuli were presented in 5 sets equated in range of stimuli used. There was no significant difference between the two groups in reliability (as represented by interjudge agreement) on the first three sets of stimuli. On the fourth and fifth sets of stimuli there was a drop in reliability of the group judging individually, causing a significant difference between the two groups on these trials. Two more random groups were tested in exactly the same way with the addition of stress instructions

that informed the Ss that those who made less reliable ratings of the stimuli were in need of psychological counseling. The results for the two stress groups paralleled that of the neutral group. Stress instructions had no significant effect on the Ss receiving them. A differential effect had been hypothesized, the expectation being that stress instructions would "take" in an individual setting but not in a group situation. In the former stress instructions would produce a decrement in performance; in the latter they would produce no difference in performance since these Ss would not believe that the instructions applied to all of them.

Perhaps some characteristic of the E affected the Ss tested in groups differently from the Ss tested individually, or the greater stress assumed by Walker et al. (in press) to be operating in the individual situation heightened the effects of fatigue for Ss tested individually. It is also quite possible that the results obtained in this study were, in fact, a chance finding.

Purpose. The present experiment compared more accurately the differential effects of group versus individual testing of HA and LA Ss on a clinical judgment task. In view of previous research and theoretical considerations, the following hypotheses are presented:

Hypothesis One. If HA judges and LA judges make many different clinical judgments individually with only the E present, the HA judges will initially be less reliable than the LA judges, but eventually will become just as reliable in their judgments as the LA judges. In the individual situation there will be sufficient stress as suggested by Walker et al. (in

press) to affect the HA and LA Ss differentially such that the HA judges will initially have stronger competing responses that will quickly be reduced to allow the HA judges to perform just as well as the LA judges.

Hypothesis Two. If HA judges and LA judges make many different clinical judgments in a non-interacting group situation, the two groups of judges will be equally reliable throughout the series of judgments. This prediction is based upon the assumption that there will be sufficiently reduced stress in the group situation such that the HA judges and LA judges will have irrelevant competing responses of comparable strength.

Hypothesis Three. The LA judges tested in the non-interacting group situation will initially be just as reliable as the LA judges in the individual situation. After making many clinical judgments the LA judges in the individual situation will become less reliable than the LA judges in the non-interacting group situation. The basis of this prediction is a frankly empirical one, as this was the finding Pribyl's (1963, unpublished) study. At the present time no theoretical explanation can be offered that will adequately explain this finding. The hypothesis is presented mainly to attempt to replicate the previously obtained results. If hypothesis three is not supported, it will imply that the previous results were due to some chance factor.

Hypothesis Four. The HA judges in the non-interacting group situation will initially be more reliable than the HA judges in the individual situation. In making many clinical judgments the HA judges in the individual situation will become just as reliable as the HA judges in the group

situation. After an even greater number of judgments, the HA judges in the individual situation will become less reliable than the HA judges in the group situation. In other words, HA judges in an individual situation are at first less reliable, then as reliable, and then again less reliable than HA judges in a group situation, assuming that a fairly large number of judgments are made.

## Method

Subjects. All of the Ss who participated in the experiment were drawn from the pool of Ss maintained at the Lake Shore Campus of Loyola University. There is a course requirement that all general psychology students must participate in five one-hour experiments. Since there are more experiments than there are subject-hours available, the students have some leeway in choosing the experiments they participate in.

As a regular classroom exercise all of the undergraduate general psychology students took the MAS during the period of time between the second and the fifth weeks of the semester. The true-false MAS items were included in a series of similar true-false items in a personality questionnaire innocuously titled the Biographical Inventory. Two graduate assistants (other than the E) in the psychology department administered the Biographical Inventory. The students were told that the Biographical Inventory was being administered in order to standardize it.

Taylor (1953) has found that there is a consistent difference in the mean MAS scores for males and females with the latter invariably scoring higher. Because of Taylor's finding and the possibility that there may be some unknown systematic difference in performance of clinical judgment tasks, only males were used in the experiment.

The male general psychology students whose scores on the MAS were in the highest 20 percent and the lowest 20 percent were selected from a group of more than 80 males who were enrolled in four of the six general psychology sections. The names of these students were put on a folder along with the



statement that they had been selected randomly for the experiment. In accord with the usual procedure for obtaining Ss, the folder (having appointment times in it) was passed around the four sections. The E tested these Ss individually.

The HA judges and LA judges for the non-interacting group condition were selected on the basis of MAS scores from the distribution of MAS scores of the 80 males in the remaining two general psychology sections held on Lake Shore Campus. As was true for the Ss in the individual condition, the HA judges were those males with MAS scores in the top 20 percent of the distribution and the LA judges were those males with MAS scores in the lowest 20 percent of the distribution. To make the group setting as natural as possible, the experiment was run during the regular class period of the two general psychology sections. Data were collected from all of the students of both sections, but only the data from the students selected on the basis of their MAS scores were analyzed.

Stimuli. The stimuli were the 100 schizophrenics' vocabulary test responses used by Hunt and Walker (1963). These stimuli had previously been standardized by experienced clinicians on a 7-point scale according to the amount of exhibited disorganization (Hunt & Jones, 1962). The stimuli were presented in 10 sets consisting of 10 responses each. Each set contained two stimuli at each of the first three scale points and one stimulus at each of the four remaining scale points of the 7-point scale used by the clinicians.

Procedure. The Ss in the individual condition were tested in one of

the experimental booths at the Lake Shore Campus. The stimuli were presented by means of a projector on a screen approximately four feet away from the S. The E sat to the left of the S and behind him, at a table on which the projector rested. The S called out his rating and the E recorded it on a data sheet like the one shown in Appendix A. After the experiment was completed the E asked each S the questions presented on the Questionnaire in Appendix B.

The Ss in the non-interacting group condition were given a data sheet like the one shown in Appendix A. They filled in their own responses. At the end of the experiment the data sheets were collected and the Questionnaire passed out. The Ss were asked to put their names on the Questionnaire (shown in Appendix B) and to fill it out.

All Ss received the same instructions. They were told that their responses would be confidential and would not influence their standing in the general psychology course. Then the following instructions, taken from Hunt and Walker (1963, p. 495) were read:

"We are going to present you with a number of responses made by schizophrenic patients to vocabulary test items taken from an intelligence test. One of the ways in which the pathology of schizophrenia may express itself is through disorganized thinking which results in atypical, unusual, or 'abnormal' responses to the items on such a test. The qualitative interpretation by the clinician of such test responses is one of the bases upon which he may make a clinical or diagnostic interpretation. The extent of the disorganization exhibited in these responses is not uniform. In some of the responses it is minimal and others it is extreme.

"You are asked to rate these responses on a 7-point scale, from 1 through 7, according to the severity of the disorganization exhibited in the response, with the low end of the scale representing minimal disorganization and the high end of the scale representing maximal disorganization. In making these ratings we are

asking you to concentrate upon the severity of the disorganization exhibited in the response. In essence, what we are asking you to do is to judge how 'schizophrenic' each response is. Some responses will seem quite normal; those you would rate '1'. Others will be so disorganized as to require a '7' rating. The majority will fall somewhere in between.

"We are now going to project onto the screen a stimulus word and the response to it. Think out your rating carefully, but as soon as you make up your mind give your response.

"First you will be given three practice trials. Do you have any questions before we begin? You will have an opportunity to ask questions after the practice trials, but once the experiment starts you will have to hold all questions until the end of the experiment. I shall be glad to answer any additional questions at that time."

Then three practice stimuli were presented. The ratings given by the clinical psychologists were announced to the Ss as the appropriate slide was presented. At this time any questions were answered.

The ten trials were then presented. In the non-interacting group condition the stimuli were presented for approximately five seconds each. Enough time was taken between sets to change the slide tray in the projector and announce the number of the next set. For the Ss run individually each slide was presented only for the amount of time that it took the S to give his rating. Between sets the E simply changed the slide tray in the projector.

After the Questionnaire had been filled out the E answered any additional questions and requested that the Ss not discuss the experiment with their friends.

## Results

After all of the data were collected, data from Ss who did not follow directions or who knew about the experiment beforehand were eliminated. After the above mentioned Ss' data were eliminated, there were four groups of 16 Ss each. The range of the MAS scores of the HA judges who performed in the non-interacting group condition was from 23 to 38 while the range of MAS scores for the HA judges performing in the individual condition was from 26 to 34. The range of the MAS scores for the LA judges tested in class was from 1 to 7 while the range for the LA judges tested individually was from 1 to 9. The ranges of scores are quite comparable to previous research in this area.

In the data analysis each S's ratings of each set of 10 stimuli were correlated with the ratings for the same set of stimuli of each of the other members of his group. Pearson's product-moment correlation coefficient was used. Each of the r's was converted into a z' value according to the table presented by Edwards (1960). Mean z' values were then computed for each of the four 16-member groups of Ss on a set-by-set basis.

Duncan's new multiple range test was used to test the significance of the differences between the means of the four groups on a set-by-set basis. The Duncan's range test was used to eliminate the spuriously large number of significant t values that would be obtained if a single mean were used in more than one comparison. The set-by-set means and the results of the range test are presented in Appendix C. The comparisons between the pairs of means that are of interest in this study are presented in Table 1 and

Table 2. In all there are 60 possible combinations of pairings of two means through all of the 10 range tests that were done. Of the 60 comparisons, there are three pairs of means that are significantly different at the .01 level and four pairs of means that are significantly different at the .05 level. Two of the pairs of means that are significantly different at the .05 level and one of the pairs of means that is significantly different at the .01 level were not predicted by the hypotheses of this study. Of the 40 pairings of means that are relevant to this study, two are significantly different at the .05 level and two at the .01 level. Since the protection level against Type I errors for the Duncan's range test where all combinations of pairings of four means are tested at the .05 level is 86 percent, the conclusion is that those differences that were found to be significant were due primarily to chance. The protection level against Type I errors for all combinations of four means at the .01 level is 97 percent. In this case, too, one must conclude that the differences found were due to chance.

Table 1

Significance of Differences Between Mean Interjudge Reliabilities  
for 16 High Anxious (HA) and 16 Low Anxious (LA)  
Ss in Group and Individual Conditions

Set	Group		Individual	
	HA Mean <sup>a</sup>	LA Mean <sup>a</sup>	HA Mean <sup>a</sup>	LA Mean <sup>a</sup>
1	1.012	1.044	0.843	1.144
2	0.764	0.756	0.510 *	0.932
3	1.062	1.257	0.997	1.256
4	1.248	1.279	0.906	1.144
5	0.849	0.821	0.685	0.873
6	0.840	0.928	0.593 **	1.042
7	0.947	1.129	0.928	1.028
8	1.281	1.272	0.829	1.109
9	0.702	0.814	0.661	0.917
10	1.126	1.162	0.896	1.194

<sup>a</sup>All means are  $z'$  values

\*Difference between means significant at .05 level according to Duncan's new multiple range test.

\*\*Difference between means significant at .01 level according to Duncan's new multiple range test.

Table 2

Significance of Differences Between Mean Interjudge Reliabilities for 16 High Anxious (HA) Ss in the Group (G) Condition and 16 HA Ss in the Individual (I) Condition and for 16 Low Anxious (LA) Ss in the G Condition and 16 LA Ss in the I Condition

Set	High Anxious		Low Anxious	
	G Mean	I Mean	G Mean	I Mean
1	1.012	0.843	1.044	1.144
2	0.764	0.510	0.756	0.932
3	1.062	0.997	1.257	1.256
4	1.248 *	0.906	1.279	1.144
5	0.849	0.685	0.821	0.873
6	0.840	0.593	0.928	1.042
7	0.947	0.928	1.129	1.028
8	1.281 **	0.829	1.272	1.109
9	0.702	0.661	0.814	0.917
10	1.126	0.896	1.162	1.194

<sup>a</sup>All means are  $z'$  values.

\*Difference between means significant at .05 level according to Duncan's new multiple range test.

\*\*Difference between means significant at .01 level according to Duncan's new multiple range test.

A 2 x 2 x 10 analysis of variance was done on the data, the variables being level of anxiety, social situation, and sets of stimuli. The results of this analysis are presented in Table 3. Values of  $F$  significant at the .01 level are marked with an asterisk. The results indicate a significant  $F$  for anxiety,  $F$  being equal to 27.409 ( $df$  1 and 60,  $p < .01$ ). Inspection of Figure 1 shows that, in general, the main effect of anxiety was due to the greater reliability of LA  $S_s$ .

The  $F$  for social situation was 7.694 ( $df$  1 and 60,  $p < .01$ ). While the relationship is complex, the reliability of  $S_s$  in the group situation is, in general, significantly greater than is the interjudge reliability in the individual situation. The interaction effect is also significant,  $F$  being 10.883 ( $df$  1 and 60,  $p < .01$ ). This indicates that the effect of the social situation on clinical judgment is not independent of the anxiety level of the subject.

The effect of the 10 sets of stimuli is significant at the .01 level,  $F$  being 53.778 ( $df$  9 and 540). This can be interpreted as being due to a position effect or an item content effect. The interaction between anxiety and sets is not significant, indicating that reliability varies uniformly over sets for both levels of anxiety.

The social situation by sets interaction is significant,  $F$  being 5.487 ( $df$  9 and 50,  $p < .01$ ). It appears that the fluctuation in interjudge reliability from set to set is not independent of the social situation.

The  $F$  for the triple interaction, which takes into account the three variables of anxiety level, social situation, and sets as influences on the



Table 3

Analysis of Variance of the  $z'$  Values for the Mean Interjudge Reliabilities for the Four Groups of Subjects

Source of Variation	Sum of Squares	df	Mean Square	F
Anxiety	4.687	1	4.687	27.409*
Social Situation	1.308	1	1.308	7.649*
Anxiety x Social Situation	1.861	1	1.961	10.883*
Error (a)	10.245	60	.171	
B Sets	15.004	9	1.667	53.778*
Anxiety x Sets	.416	9	.046	1.484
Social Situation x Sets	1.528	9	.170	5.487*
Anxiety x Social Situation x Sets	.777	9	.863	27.839*
Error (b)	16.620	540	.031	
Total	52.446	639		

\*p is at 1% or less

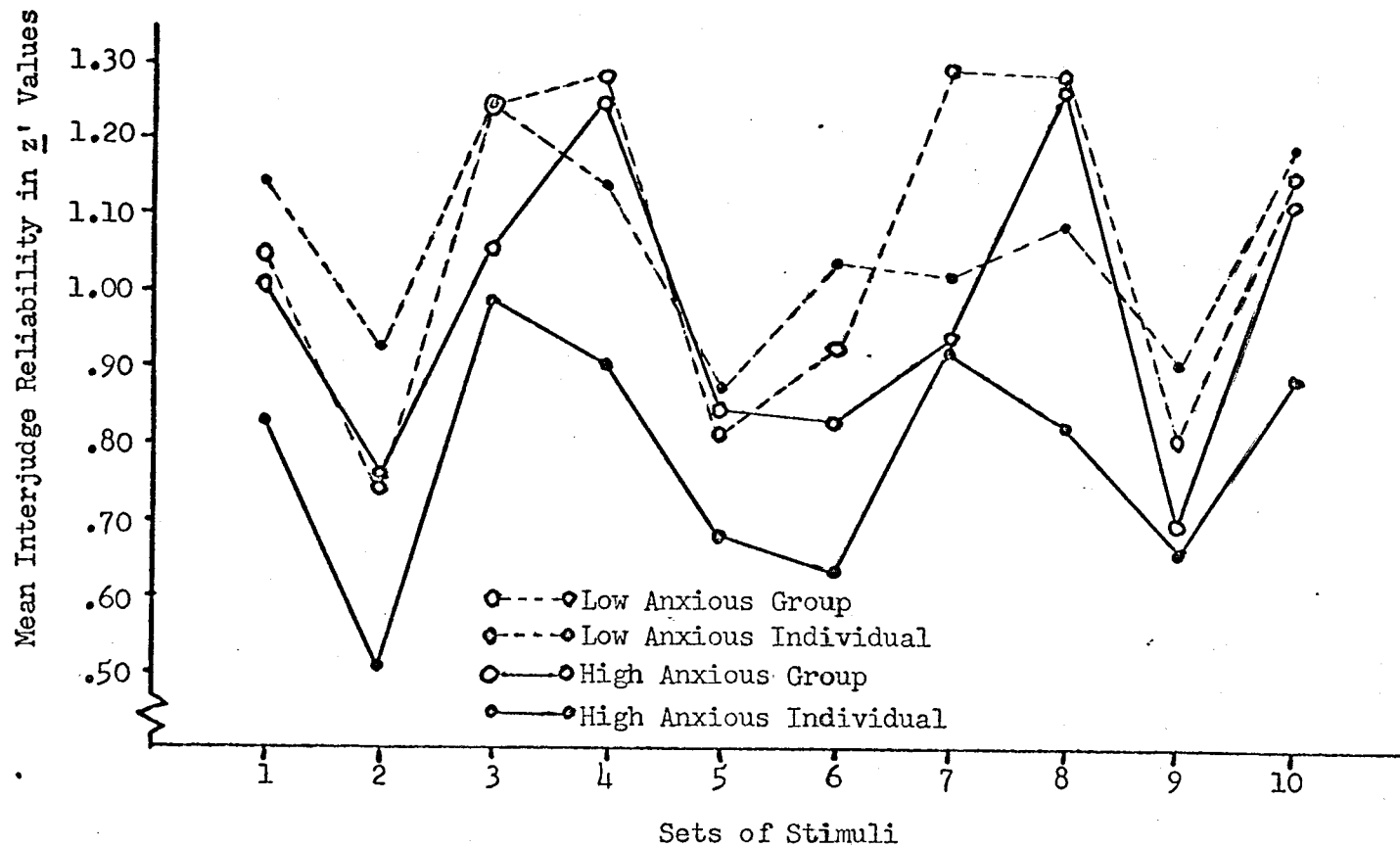


Fig. 1. Mean interjudge reliability (in  $z'$  values) on a set-by-set basis.

interjudge reliability of the subjects. This, too, is significantly larger than chance,  $F$  being 27.839 ( $df$  9 and 540,  $p < .01$ ). This result means that when these three variables exist in an experiment, they do not act independently.

The fluctuations from set to set in reliability over all judges led to a postdiction that the sets of stimuli might in themselves be of varying difficulty for the naive judges. To test this postdiction the mean standard deviations of the clinicians that originally standardized the stimuli were computed on a set-by-set basis. The clinicians' mean standard deviations from the ratings obtained in the standardization were taken as an index of difficulty. These standard deviation scores and the mean  $z'$  values for all four groups were graphed on the same grid on a set-by-set basis. This is presented in Figure 2.

If the postdiction were to be supported, the mean clinician standard deviation should be low when the naive judges' interjudge reliability is high. This relationship holds only when going from trial 4 to 5, from 6 to 7, possibly from 8 to 9, and from 9 to 10. It must be noted that in constructing the graph, the units for the two scales were not equated. However, no desirable transformation would change the order of the variables. Apparently the changes in interjudge reliability from set to set cannot be interpreted solely as a function of differential difficulty of the sets.

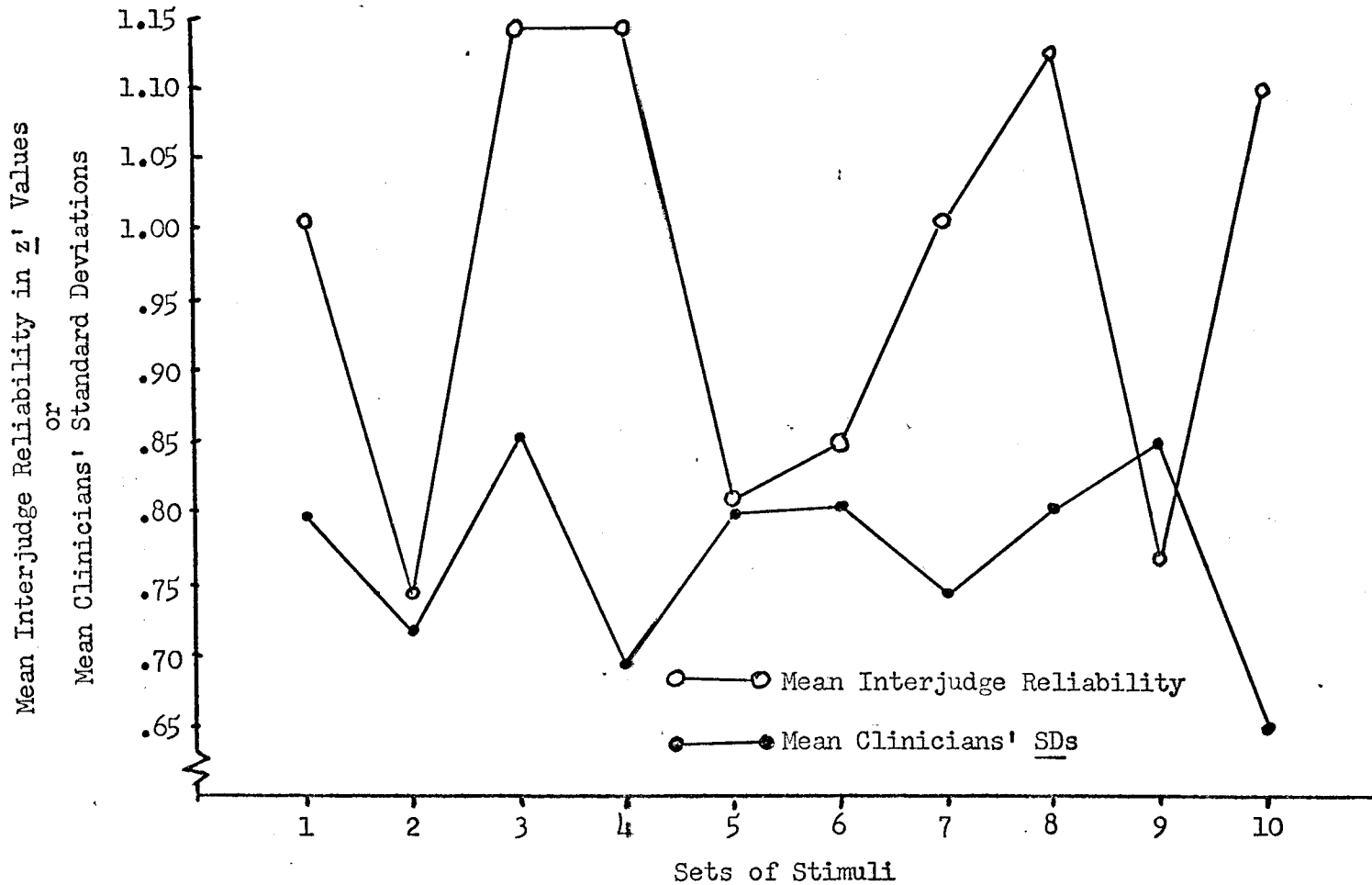


Fig. 2. Mean interjudge reliabilities (in  $\underline{z}$ ' values) of the four experimental groups combined for each set of stimuli and the mean standard deviations of the clinicians' ratings of each set of stimuli.

## Discussion

Contrary to the predictions of hypothesis one, the HA judges in the individual condition did not show an initial decrement in reliability. This result is not in agreement with the previous research of Hunt and Blumberg (1961) and Hunt and Walker (1963) who found that HA judges were less reliable than LA judges on the initial trial in three independent studies.

Because this initial decrement for HA §s was not found, the present study fails to support the Hunt and Walker (1963) hypothesis that HA judges had more of what Child (1954) called "task irrelevant responses" due to anxiety than did LA judges. Theoretically HA §s should have done worse at first and then, once the task irrelevant responses were eliminated, should have performed on a par with the LA §s.

That the expected result did not occur is surprising since the stimuli and methodology used replicated the Hunt and Walker (1963) study exactly. One can only guess that perhaps some unknown selection factor resulted in differing populations for the two studies or that some E variable influenced the results.

Since the Hunt and Walker study was done at Northwestern University and the present study at Loyola University, some unknown selection factor may have been a critical variable. Even if the two populations are similar, it is still possible that one or the other sample was biased in some unknown direction.

That an experimenter variable influenced the results is also a tenable hypothesis since the amount of stress in a given experiment could be related

to E characteristics. (Rosenthal, 1964). The second hypothesis predicted that in a group situation there would be no difference in performance of HA and LA subjects on any trials. This hypothesis was confirmed. Its theoretical relevance is, however, limited by the lack of support of hypothesis one since it adds no support to the Walker et al. (in press) assumption that the group testing condition is less stressful.

Spence (1963) suggests that the amount of task irrelevant responses is a function of the amount of stress in the experimental condition. If the assumption of the group condition being less stressful than the individual condition were correct, the differential amount of task irrelevant responses of HA and LA judges should be greater in the individual condition. The negative results of hypothesis one of this study indicate that either the assumption of differential stress for group and individual condition is invalid or that Spence's (1963) concept of task irrelevant responses being a function of stress is invalid. It is impossible to indicate from the results which is the case.

Hypothesis three predicted no difference for LA Ss on the first trial as a result of individual or group testing. As was the case with hypothesis two, this finding of no difference is not theoretically relevant since its importance depended on finding a significant difference between HA and LA Ss on the first trial of the individual condition.

It was predicted by hypotheses three and four that both HA and LA Ss in the individual situation would have a relatively poorer performance on later trials. The same variables that influenced the failure of the results

to support hypothesis one may well have been critical determinants here.

Since the Pribyl (1963) finding had no foundation in theory it is now even more probable that it was no more than a chance result. On the other hand, it should not be dismissed too lightly in view of the fact that random groups of Ss participated in Pribyl's study, while in the present study highly selected Ss (HA and LA) participated in the individual and group conditions.

In hypothesis four the HA judges in the individual condition were assumed to be in a more stressful experimental situation than the HA judges in the group condition. This was apparently not the case as there was no difference found in the initial trials of the two conditions for the HA judges. It was thought that HA judges in the individual condition would show more task irrelevant responses than HA judges in the group condition because of less stress in the latter condition. Either Spence (1963) is not correct in her assumption that task irrelevant responses are a function of the amount of stress, or there was insufficient stress in the individual condition. It is difficult to choose between these two explanations as previous research did not directly test hypotheses relating clinical judgment and the effects of testing Ss in groups and individually.

Although no predictions were made concerning the analysis of variance, these results are nevertheless interesting. One finding of importance is that the Ss in the group condition were more reliable than the Ss in the individual condition. This result lends support to the applicability of Hunt's (1959) analogy of psychophysics and clinical judgment since it agrees

with findings of Allport (1920, 1924) using judgments of sensory stimuli. He found that sensory judgments made in a co-acting but non-interacting group were less extreme than if the judgments were made alone.

The reliability measure used in this study tells essentially how well the SS agree with each other. The higher the reliability the more alike are all the SS judgments of the stimuli. Allport (1920, 1924) also made observations somewhat parallel to the results of this study; these indicate that SS in the group condition were more reliable than SS in the individual condition. He noticed that SS' free associations were more common or less idiosyncratic if they were made in a co-acting but non-interacting group. This parallel further points out the generality of phenomena that take place in clinical judgment.

When all of the LA SS were combined, they were found to be more reliable than the HA SS. One very speculative explanation for this might be that the HA SS did make some task irrelevant responses that were not dissipated as the trials progressed. This is quite possible since new stimuli were presented on every set, and it may be that HA SS made task irrelevant responses to specific stimuli as they were presented. These responses may have been small enough in number to produce non-significant results in the Duncan's test of mean differences, but their cumulative effect on the performance of HA SS could have been picked up by the more sensitive F test.

The finding that there was a good deal of variance contributed by the sets of stimuli and the finding that the amount of variance was not uniform for the group and individual conditions suggest that sets of stimuli used in



clinical judgment studies should be standardized under these conditions. Situational and individual difference variables are proving to be very important in clinical judgment and future research must take these into account.

## Summary

Previous research has shown that the relative performance of high anxious (HA) and low anxious (LA) Ss (operationally defined by extreme scores on the Taylor MAS) on the initial trials of clinical judgments is different depending whether the HA and LA Ss judged in a group or individually. In this study the Ss were given the judgmental task of rating the amount of confusion exhibited in 100 vocabulary test responses taken schizophrenics' test protocols. The Ss rated the stimuli in ten trials or sets consisting of 10 stimuli each. Because of previous research HA Ss were expected to perform less well than LA Ss on the first trial when the Ss judged individually. This expectation was not born out. When HA Ss and LA Ss worked in a group there was no difference in performance as expected, but since a differential effect due to working in a group or individually was predicted the implications of this finding are limited. Further hypotheses predicting a decrement in performance for both HA and LA Ss judging individually on later trials were presented and tested but not supported. The lack of replication of previous research was discussed in terms of differing subject populations and E variables. For exploratory purposes an analysis of variance was done on the r' values of the mean interjudge reliabilities, (essentially, a measure of how well the judges agreed with each other), the variables being level of anxiety (HA or LA), social situation (group or individual condition), and sets of stimuli. It revealed that clinical judgments, like sensory judgments, tend to be more alike (better interjudge agreement) if the Ss judge in a group than if the Ss judge individually.

This latter finding suggests that many of the characteristics of clinical judgment may be similar to those found in other types of judgment.

## References

- Allport, F.H. The influence of the group upon association and thought. J. exp. Psychol., 1920, 3, 159-182.
- Allport, F.H. Social psychology. Boston: Houghton-Mifflin, 1924.
- Beebe-Center, J.G. The law of affective equilibrium. Amer. J. Psychol., 1929, 41, 54-69.
- Blumberg, Stanley. Performance changes in repeated judgments of schizophrenic verbal responses. J. clin. Psychol., 1961, 17, 308-311.
- Campbell, D.T., Hunt, W.A., & Lewis, Nan A. The effects of assimilation and contrast in judgments of clinical materials. Amer. J. Psychol., 1957, 70, 347-360.
- Campbell, D.T., Lewis, Nan A., & Hunt, W.A. Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored. J. exp. Psychol., 1958, 55, 220-228.
- Child, I.L. Personality. Ann. Rev. Psychol., 1954, 5, 149-170.
- Cline, V.B. Ability to judge personality assessed with a stress interview and sound film technique. J. abnorm. soc. Psychol., 1955, 50, 183-187.
- Edwards, A.L. Experimental design in psychological research. (rev. ed.) New York: Holt, 1960.
- Gibson, Eleanor J. Improvement in perceptual judgments as a function of controlled practice or training. Psychol. Bull., 1953, 50, 401-431.
- Grigg, A.E. Experience of clinicians, and speech characteristics and statements of clients as variables in clinical judgment. J. consult. Psychol., 1958, 22, 315-319.
- Helson, H. Adaptation level as frame of reference for prediction of psychological data. Amer. J. Psychol., 1947, 60, 1-29.
- Helson, H. Adaptation level as a basis for a quantitative theory of frame of reference. Psychol. Rev., 1948, 55, 297-313.
- Hull, C.L. Principles of behavior. New York: D. Appleton-Century, 1943.

- Hunt, W.A. Anchoring effects in judgment. Amer. J. Psychol., 1941, 54, 395-403.
- Hunt, W.A. An actuarial approach to clinical judgment. In Bass, B. & Berg, I.A. (Eds.) Objective approaches to personality assessment. New York: Van Nostrand, 1959. Pp. 169-191.
- Hunt, W.A., & Arnhoff, F.N. Some standardized scales for disorganization in schizophrenic thinking. J. consult. Psychol., 1955, 19, 171-174.
- Hunt, W.A., & Blumberg, Stanley. Manifest anxiety and clinical judgment. J. clin. Psychol., 1961, 17, 8-11.
- Hunt, W.A., Jones, N.E., & Hunt, Edna B. Reliability of clinical judgment as a function of clinical experience. J. Clin. Psychol., 1957, 13, 377-378.
- Hunt, W.A., & Jones, N.F. The experimental investigation of clinical judgment. In Bachach, A., (Ed.) Experimental foundations of clinical psychology. New York: Basic Books, 1962. Pp. 26-51.
- Hunt, W.A., & Volkman, J. The anchoring of an affective scale. Amer. J. Psychol., 1937, 49, 88-92.
- Hunt, W.A., & Walker, R.E. Manifest anxiety and clinical judgment--A re-examination. J. clin. Psychol., 1963, 29, 494-497.
- Hoffman, P.J. The paramorphic representation of clinical judgment. Psychol. Bull., 1960, 57, 116-131.
- Johnson, D.M. The psychology of thought and judgment. New York: Harper, 1955.
- Jones, N.F. Context effects in judgment as a function of experience. J. clin. Psychol., 1957, 13, 379-382.
- Meehl, P.E. Clinical vs. statistical prediction. Minneapolis: Univ. Minnesota Press, 1954.
- Pribyl, J.F. A comparison of the reliability of clinical judgments made in a group and individually under stress instructions, unpublished.
- Rosenthal, Robert. Experimenter outcome-orientation and the results of the psychological experiment. Psychol. Bull., 1964, 61, 405-412.
- Sarason, I.G. Empirical findings and theoretical problems in the use of anxiety scales. Psychol. Bull., 1960, 57, 403-415.

- Sarbin, T.R., Taft, R., & Bailey, D.E. Clinical inference and cognitive theory. New York: Holt, 1960.
- Spence, Janet T. Learning theory and personality. In Wepman, J.M. & Heine, R.W., (Eds.) Concepts of personality. Chicago: Aldine, 1963, Pp. 3-30.
- Taylor, Janet A. The relationship of anxiety to the conditioned eyelid response. J. exp. Psychol., 1951, 41, 81-92.
- Taylor, Janet A. A personality scale of manifest anxiety. J. abnorm. soc. Psychol., 1953, 48, 285-290.
- Taylor, Janet A. Drive theory and manifest anxiety. Psychol. Bull., 1956, 53, 303-320.
- Underwood, B.J. Psychological research. New York: Appleton-Century-Crofts, 1957.
- Walker, R.E., Hunt, W.A., & Schwartz, M.L. Manifest anxiety and clinical judgment in a group setting. J. soc. Psychol., in press.

**Appendix A**

Name \_\_\_\_\_

Trial I	Trial II	Trial III	Trial IV	Trial V
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10

Trial VI	Trial VII	Trial VIII	Trial IX	Trial X
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10



**Appendix B**

Questionnaire

1. Did you know anything about this experiment beforehand? Yes \_\_\_\_\_  
No \_\_\_\_\_

2. If you did know anything about this experiment beforehand, what did you know about it? \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

3. Did you understand what you were expected to do? Yes \_\_\_\_\_ No \_\_\_\_\_

4. If the answer to the previous question was no, what didn't you understand? \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

5. Please comment below on the experiment or any of the above questions.

**Appendix C**

**Duncan's New Multiple Range Test Applied to the Differences  
Provided by All Combinations of the Mean Interjudge  
Reliabilities for All Four Experimental  
Groups on a Set-by-Set Basis**

Set	Means			
1	HAI <sup>a</sup> 0.843	HAG <sup>b</sup> 1.012	LAG <sup>c</sup> 1.044	LAI <sup>d</sup> 1.144
2	HAI 0.510	LAG <u>0.756</u>	HAG 0.764	LAI 0.932
3	HAI 0.997	HAG 1.062	LAI 1.256	LAG 1.257
4	HAI 0.906	LAI <u>1.144</u>	HAG 1.248	LAG 1.279
5	HAI 0.685	LAG 0.821	HAG 0.849	LAI 0.873
6	HAI 0.593	HAG <u>0.840</u>	LAG 0.928	LAI 1.042
7	HAI 0.928	HAG 0.947	LAI 1.028	LAG 1.129
8	HAI 0.829	LAI <u>1.109</u>	LAG 1.272	HAG 1.281
9	HAI 0.661	HAG 0.702	LAG 0.814	LAI 0.917
10	HAI 0.896	HAG 1.126	LAG 1.162	LAI 1.194

Note.--All means are  $\bar{z}$ ' values. Any two means not underlined by the same line are significantly different. Any two means underlined by the same line are not significantly different.

<sup>a</sup>High Anxious Individual (N=16)

<sup>b</sup>High Anxious Group (N=16)

<sup>c</sup>Low Anxious Group (N=16)

<sup>d</sup>Low Anxious Individual (N=16)

\*Significant at .05 level

\*\*Significant at .01 level

APPROVAL SHEET

The thesis submitted by Joseph F. Pribyl has been read and approved by three members of the Department of Psychology.

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 1, 1962  
Date

Ronald E. Bridges, Ph.D.  
Signature of Adviser