GATING:

THE APPLICATION OF A PERCEPTUAL THEORY TO THE ISSUE OF DIAGNOSIS IN PSYCHOTHERAPY

ROBERT BERNARD LEVINSON

A THESIS FOR DEGREE OF PH D
MICHIGAN STATE UNIVERSITY

1957

This is to certify that the

thesis entitled

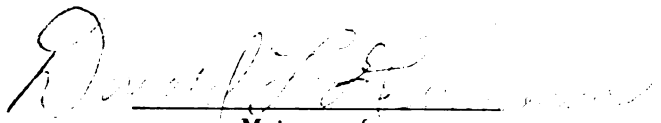
THE APPLICATION OF A PERCEPTUAL THEORY
TO THE ISSUE OF DIAGNOSIS IN PSYCHOTHERAPY

presented by

ROBERT B. LEVINSON

has been accepted towards fulfillment
of the requirements for

PH.D. degree in PSYCHOLOGY



Major professor

Date MARCH 7, 1958

O-169

G A T I N G :

THE APPLICATION OF A PERCEPTUAL THEORY TO THE ISSUE OF DIAGNOSIS IN PSYCHOTHERAPY

BY

ROBERT BERNARD LEVINSON

A THESIS

Submitted to the School for Advanced Graduate Studies of Michigan State University of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1957

"And once set up, a system probably does as much harm as it does good. It serves as a sort of sacred grating behind which each novice is commanded to kneel in order that he may never see the real world, save through its interstices."

Edward C. Tolman

ACKNOWLEDGEMENTS

In the course of pursuing a doctoral degree an individual becomes indebted to many people for the assistance offered to him. At this time, the author of this dissertation would like to acknowledge the help he has received.

For the guidance in helping the author steer through the sometimes troubled waters of advanced Psychology courses and for serving as the original chairman of this doctoral committee, before going on a sabbatical leave, many thanks are due to Dr. Donald M. Johnson.

To Dr. Donald L. Grummon, who graciously accepted the role of chairman of this doctoral committee, and who gave generously of his time and effort to the development of this investigation, the author wishes to express his gratitude.

The author would also like to express his appreciation to the other members of his doctoral committee: Drs. M. Ray Denny, Bill L. Kell, G. Marion Kinget, and Louis L. McQuitty.

A debt of gratitude is owed to the Veterans Administration for making it possible for the author to conduct this research by providing not only many of the clinicians who served as Judges, but also by making time available for the collection of data.

This study could not have been conducted without the cooperation of the clinicians who served as Judges and Mr. "L" who was the client. To the Judges who are listed in Appendix A and the client, Mr. "L", the author offers his many thanks.

Finally, to the two women who made this whole project seem worthwhile, this work is dedicated:

"Thanks Betty."

"Mindy, maybe you'll get to know your Daddy now."

AN ABSTRACT

Gating: The Application of a Perceptual Theory to the Issue of Diagnosis in Psychotherapy

by

Robert Bernard Levinson

Gating is a process introduced by Bruner (1) as part of a perceptual theory. It is a mechanism which permits the passage of certain stimuli in the environment (those in accord with an individual's category system) and blocks passage of other stimuli (not in accord with the category system). A category system may be conceived of as a network of sets, or personal constructs, which evolve from a process of learning how to isolate, weigh, and use critical attributes or cues, and which is markedly influenced by individual expectations and needs.

The process of gating has direct implications concerning the use of formal diagnostic procedures in psychotherapy. The commonly accepted view -- "diagnostic" position -- holds that pre-therapy plans (based on diagnostic techniques) enhance a therapist's understanding and, thereby, increase his effectiveness with a client. Adherents of an "a-diagnostic" viewpoint, however, suggest that pre-therapy formulations may blunt the

therapist's understanding of his client. That is, the operation of a gating mechanism could result in a selective screening out of certain of the stimuli present in therapeutic interviews--namely, those cues not in accord with the pre-therapy schema--while, at the same time, other cues (more in accord with the pre-therapy plan) may be over-evaluated.

This study concerned itself with the following general proposition:

> the type of preconceived categorization system with which one approaches events will affect the sensitivity to nuances in those events

The proposition was investigated through two, more specific, hypotheses.

Hypothesis 1:

> the sensitivity of an individual to nuances of a situation will be diminished if he approaches that situation with a preconceived categorization system

Hypothesis 2:

> the sensitivity of one individual for another person will increase as the degree of congruence between the categorization systems of the two individuals increases

Using a Q-methodology design, 32 clinical psychologists, (in four groups of eight each) after doing a self-sort, attempted to predict a client's self-concept (as revealed by Q-sorts) on the basis of diagnostic and interview material. Each Judge made two predictions. The dependent variables were the accuracy scores--the degree to which the predictions approximated the client's actual sortings. The independent

variable was the sequence in which the stimulus material was presented. Four sequences were employed: (i) Interview (I") material before Diagnostic (D) material; (ii) D I"; (iii) I' I"; (iv) I" I"--conditions (iii) and (iv) represent controls; I' and I" being, respectively, the 11th and 13th tape recorded therapy sessions.

Using an analysis of variance technique, the results obtained indicate:

> Diagnostic material does not markedly enhance nor significantly decrease the sensitivity of a Judge to Interview material; and, similarly, for the effect of Interview material on Diagnostic material
>
> Although the presentation sequence of the stimulus material does not significantly affect sensitivity, making two predictions, particularly if based on widely differing types of material, results in significantly less accuracy for the second prediction
>
> Support was found for the validity of Hypothesis 1
>
> The validity of Hypothesis 2 remains in question; the results were in the anticipated direction but did not reach the level of statistical significance
>
> The general proposition underlying this study was, in the main, supported by the findings

The limitations of this investigation, as well as its implications for clinical practice and future research, were discussed.

---

(1) Bruner, J.S. "On perceptual readiness" Psych. Rev. 1957 64:123-152

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

*****
***
*

INTRODUCTION

It is the intention of this study to bring some empirical evidence to bear on the issue regarding the advisability of pre-therapy diagnostic formulations. In view of the amount of time spent in doing diagnostic "work-ups" for therapy candidates (cf.87,p.101-2) as well as the amount of heat generated in discussions between clinicians of differing persuasions regarding the usefulness of such reports, it was felt that an attempt to bring some light into this controversy would be a worthwhile endeavor.

Further, this study will also serve as an empirical test for an hypothesis translated from the area of perception into the realm of clinical psychology. Critics, as well as clinicians themselves, often bemoan the lack of well formulated theories in the clinical field. Perhaps studies involving such translations can lead to a rapprochment between the more theoretically oriented areas in psychology and clinical psychology.

To be more explicit, this investigation is concerned with the general hypothesis that an individual's sensitivity to the nuances present in a situation is markedly

affected by the type of preconceived categorization system with which he approaches that situation. That is, if, say, a clinician were to enter into a therapeutic relationship with pre-formed "notions" about the client (based on, say, diagnostic protocols) such a therapist might be less sensitive to certain cues in the therapy session than if he had not had this presumptive formulation. The focus of this investigation is the effect that preconceived ideas have on sensitivity.

In what follows the theoretical background and empirical research pertinent to the above area of interest will be reviewed. The definitions and assumptions underlying the general and specific hypotheses, as well as the rationale and procedures for testing these hypotheses, will be detailed. A discussion of the results obtained and their implications for future research will follow. The final chapter will contain a summary and the conclusions of this investigation.

BACKGROUND OF THEORY AND RESEARCH

Diagnosis and Psychotherapy

Psychodiagnostic procedures appear to have earned a dual role in relation to psychotherapy. The first use made of diagnostic formulations is to provide a foundation for devising a plan (1,71,106) or establishing a rational (99) and scientific (89) basis for the choice of treatment (10). In its second role diagnostic techniques serve as adjunctive procedures in therapy. In this latter function they are employed to: direct therapeutic activity toward fundamental problem areas (62,72); establish rapport (5,21); overcome plateaus (62,72) or resistance to interpretations (5) in the therapy sessions; as a means of becoming better acquainted with the dynamics of the client's thoughts (5); and, to supply information for use in future interviews.[1]

In both these roles, moreover, there is a close interaction between the two fields of clinical interest. In Watson's words (106,p.26) "If diagnosis does not imply treatment, the diagnostic study would be of little value."

---

[1] A possible third role for diagnostic procedures might be delineated--research. Muench (73), for example, cites the following ways in which diagnosis has been oriented to therapy: to determine treatability, to determine success, to study the process of therapy, and to make comparative evaluations of therapeutic techniques.

The most commonly accepted generalization made from this diagnostic-therapeutic interaction is that it has resulted in cross-fertilization of both fields. For example, the insights achieved in therapeutic interviews have lead to a verification and enrichment of diagnostic interpretations (e.g.,90).

Despite the wide acceptance of this mutual-enhancement viewpoint, there is not universal agreement among clinicians that such is actually the case. To some practicing psychotherapists the advisability of making the results of diagnostic tests known to the therapist who will see the client is a point at issue. While the adherents of the more traditional "diagnostic" point of view contend: "The diagnosis...is an essential part of the approach to the case" (15); proponents of the opposing or "a-diagnostic" viewpoint maintain: "Diagnostic knowledge and skill is not necessary for good therapy...." (88;cf.87,p.101-2).

Thus there are two conflicting positions on this issue. The rationale for each of these points of view will be examined more closely in the following sections.

Diagnostic Viewpoint

Diagnosis is conceived of as a formal description of a client and his behavior by a variety of methods whose basic purpose is to discover the personality dynamics of each individual case. It is implied that the more complete

the description, the more complete will be the therapist's understanding of why and how the individual got that way. This, in turn, will enhance the therapist's effectiveness, and, hence, the effectiveness of the entire therapeutic undertaking: "Diagnostic constructs should sensitize the clinician to respond to significant characteristics of the client's behavior that might otherwise have been overlooked" (10).

In support of this thesis Thorne (100) lists the following as among the important objectives of diagnosis:

1. to demonstrate the aetiological factors
2. to differentiate between organic and functional disorders
3. to discover the personality reaction of the organism to its disability
4. to discover the extent of organic change with resulting functional disability
5. to estimate the extensity or intensity of the morbid process in relation to actuarial data concerning type and severity
6. to determine the prognosis or probable course
7. to provide a rational basis for specific psychotherapy
8. to provide a rational basis for discussing the case with the patient and relatives
9. to provide a scientific basis for classification and statistical analysis of data

In Thorne's words: "...the experienced clinician utilizes diagnosis as the foundation for all rational case handling" (100).

A-diagnostic Viewpoint

Opposed to the traditional orientation, outlined above, are those clinicians who see diagnostic formulations as possible obstacles to an effective therapeutic relationship. Perhaps the most vocal group holding this "a-diagnostic" view are those therapists of the Client-Centered persuasion. Rogers (88) writes:

> The therapist must lay aside his preoccupation with diagnosis and his diagnostic shrewdness, must discard his tendency to make professional evaluations, must cease his endeavors to formulate an accurate prognosis, must give up the temptation subtly to guide the individual, and must concentrate on one purpose only; that of providing deep understanding, and acceptance of the attitude consciously held at this moment by the client as he explores step by step into the dangerous areas which he has been denying to consciousness.

The rationale offered for this point of view cites two basic objections to the "diagnostic" approach:

> ...the very process of psychological diagnosis places the locus of evaluation so definitely in the expert that it may increase any dependent tendencies in the client.... When the client perceives the locus of judgment and responsibility as clearly resting in the hands of the clinician, he is, in our judgment, further from therapeutic progress than when he came in. (84,p.223)

> ...it has certain social and philosophical implications which need to be carefully considered and which, to the writer, are undesirable. ...the long-range social implications are in the direction of the social control of the many by the few. (84,p.224)

In addition Rogers maintains that diagnostic formulations may blunt rather than enhance the therapist's understanding of what the client experiences during the therapy sessions.

Other objections, based on empirical evidence and relevant to several of Thorne's points (p.5) can be cited. Powdermaker (80) notes that the diagnostic approach leads to "a gathering of data which are not obviously and immediately concerned with the patient's problem;" (cf. Thorne's point 3).

Soskin (93) found both experienced and novice clinicians, when using the Rorschach and/or the Thematic Apperception Test as a basis for conceptualizing a subject, appear to be predisposed to anticipate more maladjustive tendencies; (cf. Thorne's point 5).

Patterson (77) questions the usefulness of diagnoses in selecting a therapeutic approach "since the choice depends more on the specific training, experience, and preferences of the therapist than upon the diagnosis;" (cf. Thorne's point 7). Further, Herrick & Nagy (56) found a correlation of .15 between the ability to make a diagnosis and the ability to choose proper remedial techniques.

Ash (3) found the reliability of psychiatric diagnoses to be .20. Mehlman's (70) more recent study concluded: "The existing system of psychiatric classification can probably

have little value for administrative management of patients, or for research...." (cf. Thorne's point 9).

Gill et al (49) refer to diagnostic procedures as characteristic of "the older psychiatric textbooks;" and, Coleman (16) contends: "Nothing of importance except the patient as such and the problem of his treatment is left out in the diagnostic approach."

Thus are the lines clearly drawn which separate the opposing views regarding the usefulness of diagnostic formulations in psychotherapy.

Perceptual Theory

An empirical method, capable of evaluating the validity of the above viewpoints, was suggested by Bruner's (11) recent theoretical paper on perception. His formulations serve as the theoretical framework for this investigation.

In discussing perceptual readiness Bruner offers the thesis that perception involves an act of categorization.[2] He conceptualizes the perceptual process as follows:

> Put in terms of the antecedent and subsequent conditions from which we make out inferences, we stimulate an organism with some appropriate input and he responds by referring the input to some class of things or events. ...On the basis of certain defining or critical attributes in the input,

[2] "To categorize is to render discriminably different things equivalent, to group the objects and events and people around us into classes, and to respond to them in terms of their class membership rather than their uniqueness." (12,p.1)

> what are usually called cues....there is a selective placing of the input in one category of identity rather than another. ...A theory of perception...needs a mechanism capable of inference and categorizing.... ...all perceptual experience is necessarily the end product of a categorization process. (11,p.123)

Two reasons are offered for this position: first, the perceived event achieves its meaning from a class of percepts with which it is grouped--categorizing is equivalent to giving identity to a perceived event. The categorization system that an individual develops evolves from a process of learning how to isolate, weigh, and use critical attributes or cues for grouping events in equivalence classes. This view regarding the use made of cues is similar to Brunswik's probabilistic theory of perception. (14)

A second feature of perception is, what is perceived is "varyingly veridical"--what is perceived is somehow a representation of the external world. Bruner suggests that veridicality is not so much a matter of representation as it is "model building" (cf.105,p.185-6).

> In learning to perceive, we are learning the relations that exist between properties of objects and events that we encounter, learning appropriate categories and category systems, learning to predict and to check what goes with what. (11,p.126)

Bruner contends that a fruitful way of thinking about perceptual readiness is in terms of the accessibility of categories for use in coding or identifying events in the

environment. The likelihood that a sensory input will be categorized in terms of a given category is not only a matter of fit between sensory input and category specification, it depends also on the accessibility of a category. Accessibility is defined in terms of measures:

> The greater the accessibility of a category (a) the less the input necessary for categorization to occur in terms of this category, (b) the wider the range of input characteristics that will be "accepted" as fitting the category in question, (c) the more likely that categories that provide a better or equally good fit for the input will be masked. (11,p.129)

The relative accessibility of a category depends on the expectations of an individual regarding the likelihood of events to be encountered in his environment, and, upon the search requirements imposed on the individual by his needs. Bruner sees the categorization process involved in perception as a four step decision sequence.

1) Primitive categorization--"the silent process [Helmholtz's "unconscious inference"] that results in the perceptual isolation of an object or an event...."

2) Cue search--"a second process of more precise placement based on additional cues which may be equally silent or 'unconscious' "

3) Confirmation check--"a tentative placement of identity having occurred, the search is narrowed for additional confirmatory cues to check this placement. ...We shall speak of a selective 'gating' process coming into operation in this stage, having the effect of reducing the effective input of stimulation not relevant to the confirmatory process."

4) Confirmation-completion--"termination of cue seeking...the openess to additional cues is either normalized or 'gated out' " (11, p.130-1)

Beginning in step three, Bruner has introduced the central theoretical concept for the present study--gating. Gating is seen as a mechanism which permits passage of certain stimuli in the environment (those in accord with the subject's category system) and blocks passage of other stimuli (those not in accord with the category system). This is one of four mechanisms which mediate perception and perceptual readiness--(i) grouping and integration; (ii) access ordering; (iii) match-mismatch signaling; and, (iv) gating.

In support of his contention regarding the existance of some type of gating mechanism, Bruner cites recent physiological findings related to neural feed-back systems. That is, not only does innervation of a sensory receptor lead to a muscle response (as in the reflex arc response) but the activity of the muscle changes, to some degree, the status of the receptor. Such a sequence of events can result in the receptor becoming sensitized or de-sensitized. In the former instance it would result in an increased readiness to respond--perceptual vigilence--in the latter instance the decreased readiness to respond could be termed perceptual defense. Thus a filtering or gating system does not require the machinations of a master-mind homunculus. Rather per-

ceptual unreadiness is due to interference:

> ...failure to perceive is most often not a lack of perceiving but a matter of interference with perceiving. ...the interference comes from categorizations in highly accessible categories that serve to block alternative categorizations in low accessible categories. [11,p.145 cf.(51)]

When a subject is presented with a red four of clubs tachistoscopically (13) the tendency is to report a four of diamonds or a four of clubs with the color-suit relationship rectified. That is, the accessibility of the category "a red four of clubs" is very low, while the accessibility of the category "a red four of diamonds" or "a black four of clubs" is much higher. The reported results indicate that the input is categorized in terms of the more accessible categories, despite the perceptual distortion that this may entail.

It is this "degree of openess or closedness"--greater or lesser accessibility of categories to sensory input--which is mediated by the gating process. Stimuli appropriate for closed or low-accessible categories are "blocked" or "gated out," as in the experiment cited above, while stimuli for open or high-accessible categories are "gated in".

The similarity between "gating" and the concepts of set and attention is striking. Set, which may be defined as a readiness to make a specified response to a specified stimulus (57,p.65), also has both a positive and negative

aspect; to be set for one kind of response or one kind of stimulus necessitates an inhibition or inattention to other events in the environment. The universality of the concept of set is suggested by Allport:

> ..."being set" for a certain stimulus or for a reaction to it, and the effect which this setting has upon the ensuing or corresponding behavior are phenomena that are practically as universal as behavior itself. (2,p.211)

Johnson (57,p.67) points out that sets may be established by motives (89), social interaction (47), past experience (92), or may be induced by the task the subject is performing (66). It is suggested that the accessibility of an individual's categorization system can be "gated" in similar fashion. For example, Jones & de Charms (58) report that different inferences will be drawn from the same behavior as a function of the set promoted by the structure of the social situation.

In view of these considerations, the following derivations appear warranted. Individuals approach environmental events with a "built-in" categorization system or a network of sets; (Kelly (60,p.8) would speak in terms of "transparent templates" made of personal constructs). Such a system serves to filter out, to a greater or lesser extent--gating--those aspects of the input stimulus which are not in accord with this system. In other words, stimulus elements from an event which fit pre-existing categories will be quickly

perceived--the principle of prior entry[3]--whereas, other stimulus elements, not fitting the individual's categorization system as well, will be perceived less readily, and may even be filtered out completely; i.e., perceptual defense.

There is some experimental work extant on perception which lends support to Bruner's position concerning the gating phenomenon. For example, Wyatt & Campbell (109) have shown that if a subject develops a wrong hypothesis about the nature of a stimulus presented under sub-optimal conditions, the perception of the stimulus in terms of its conventional identity is slowed down--there is "a reduction in the adequacy of perception due to previous unverified hypothesizing or guessing." Postman & Bruner (79) have also demonstrated the detrimental effects of premature hypothesizing on veridical perception.

Theoretical Implications

A formulation such as this has direct implications for the issue concerning the use of diagnostic procedures in therapy. When diagnostic considerations have lead to the development of a plan for therapy--i.e., the client's dynamic structure, as well as his major areas of conflict have been delineated through diagnostic procedures, and

---

3 When two stimuli occur simultaneously, the one to which the individual is prepared to attend is experienced first (2,p.215).

strategies have been formulated to cope with these factors in order to attain certain specified goals--in such a situation, it may be said that the clinician approaches the therapy with a preconceived category system.

The commonly accepted view is that this plan provides additional understanding and consequently increases the effectiveness of the therapist. However, as Bruner suggests is the case when dealing with perceptual stimuli, it may be that as a consequence of the diagnostic procedures certain categories--namely, those which are in accord with the plan for the therapy--are more accessible to stimuli in the therapeutic sessions than are other categories, not in accord with this schema. It is as though the therapist is selectively attuned to receive certain types of information and to screen out other types. In effect, the therapist has become less sensitive to the entire range of stimuli.

Seen in this light, deductions from Bruner's theoretical position would suggest that therapy based on pre-conceived plans could make a therapist less effective, insofar as being aware of the total range of nuances in therapy interviews is considered a positive factor in such a relationship[4].

---

[4] The validity of this statement rests on the assumptions one is willing to make; cf. discussion in Limitations, p.61.

HYPOTHESES TO BE TESTED

This investigation is concerned with confirming or rejecting two major hypotheses which relate to the following general proposition:

> the type of preconceived categorization system with which one approaches events will affect the sensitivity to nuances in those events

A categorization system will be termed "preconceived" when it is introduced into a new situation by an individual, as a result of his experiences in other situations.

Sensitivity, as used in this study, is operationally defined as the ability to predict how another individual would describe himself on a number of selected personality traits. This definition is similar to that used in studies of empathy (7,24,25,26,52,53) and social perception (34,35, 43,44,45,67). The rationale for this definition is suggested by Luft (67):

> The judge...must take into consideration the possibility that the subject may cover up or exaggerate or falsify his "true" responses. In other words, he must be cognizant of the nature of the subject's ego-needs and his defenses, as well as the degree of insight which the subject possesses. But it is just these influences which the observer is attempting to discover in studying a subject. It therefore follows that the better the understanding, the larger will be the number of items which the judge correctly predicts.

Q-sort methodology will be employed in this study. The measure of a Judge's sensitivity will be termed his Accuracy score (ACC); the higher the ACC the more sensitivity possessed by the Judge (J)--i.e., the closer a J can predict the actual self-sort of a client (Other or "O"), the more sensitive will that Judge be deemed.

Hypothesis 1

> the sensitivity of an individual to nuances of a situation will be diminished if he approaches that situation with a preconceived categorization system

In this investigation a Judge will be asked to do a predictive Q-sort[1] based on, say, diagnostic material pertaining to an O. He will then do a second predictive-sort based on, say, interview material of this same person. Hypothesis 1 states that the "interview-sort" will be less accurate when it has been preceeded by the "diagnostic-sort" than if it (the interview-sort) had been completed first.

Rationale for Hypothesis 1

When a Judge does a predictive Q-sort he is performing a difficult--"complex and challenging"(13)--task. Other studies using this methodology find correlations between the predicted and the O's actual self-sort vary widely. For example: predictions based on hearing two successive tape re-

[1] A predictive Q-sort involves sorting a population of Q-items, by J, as he thinks O would sort the same items.

corded therapy interviews (59) correlate from -.23 to +.37 with the actual self-sort done by O; highly trained therapists attempting to predict the self-sorts of their own clients (26) have correlations from +.05 to +.85, with a median of +.41; a Judge attempting to predict an O's self-sort on the basis of diagnostic material (86) had correlations from .00 before therapy to +.55 at time of follow-up, 12 months after the therapy terminated.

As a consequence of the difficulty of the prediction task, J, in an attempt to achieve some degree of success, must put himself in O's "shoes", so to speak--an "imaginative transposing of oneself into the thinking, feeling, and activity of another" (25). This process consists of, in part, forming hypotheses (66) based on the available material as to how O would respond to each Q-item. In this fashion the Judge builds up a category system or set concerning the Q-sorting behavior of O.

The rationale, up to this point, is in accord with a more general statement of the Ames, Cantril, et al transactional-probabilistic school of thought in perception:

> ...the organism has built up certain assumptions about the world in which it lives.... All the presently existing assumptions of an individual taken together, constitute what may be called his "assumptive world". It is the only world he knows, and it determines the way he perceives the (physicalistic) world at any particular time. (2,p.278-9)

In terms of the present study, a Judge builds up a system of hypotheses or assumptions about O based on the type of material presented to him. This "assumptive world" determines in large measure the predictive sorting of the J. After doing his first predictive Q-sort, the Judge is asked to do a second predictive sort for the same O, based on different material. He approaches this second task with a category system already formed. In other words, the Judge arrives at some "notions" as to what type of a person O is on the basis of the material presented initially. Thus, he approaches the material on which he bases his second prediction with a preconceived categorization system.

The above contention assumes that categorization systems carry over from one clinical situation to another. Empirical support for this assumption is provided by Quinn (81), who found that therapists' attitudes are quite stable from client to client and are not affected by differences in the content of an hour or differences between clients (cf.74,p.1).

In accordance with Bruner's gating principle, in the situation cited above, the Judge in the second prediction session would be more open to cues or stimulus elements contained in the stimulus material which were in agreement with his preconceived notions; cues which did not fit his schema as well would be filtered out. Consequently, the Judge would not be making full use of all the cues available in the

second stimulus material. In effect, he is denying himself the use of negative evidence--negative in the sense that it runs counter to his preconceived categorization system. Therefore, the Judge's accuracy on this second prediction, as Hypothesis 1 suggests, should be less than if he had made this (second) predictive Q-sort without having completed one just prior to it.[2]

Experiments in perception (8,13,55,78,109) tend to support the rationale outlined above. It has been found that once a stimulus has been categorized in a high probability, good-fit category, the threshold for recognizing cues contrary to this categorization increases by "almost an order of magnitude" (11,p.131).

Certain types of preconceived category stystems, however, could result in increased sensitivity--namely, a categorization system which was open to an "appropriate", rather than an inappropriate segment of the total range of stimuli in a particular situation. Such a system would be one which was properly attuned to the situation at hand:

> ...the most appropriate pattern of readiness at any given moment would be that one which would lead on the average to the most "veri-

[2] There is a tacit assumption in this statement; namely, that the more "open to awareness" to all the cues in a situation the Judge is, the more sensitive he will be; cf. Rogers' proposition XVIII (84,p.520) and his discussion of therapeutic genuineness (87,p.97-8).

> dical" guess about the nature of the world around one at the moment.... And it follows from this that the most ready perceiver would then have the best chances of estimating situations most adequately and planning accordingly. (11,p.130)

Reasoning "backwards" for a moment, if an individual was able to demonstrate a high degree of sensitivity to a situation (in our terms, was able to predict the self-sort of O with a high degree of accuracy) then it could be said that he had an "appropriate" categorization system for that situation. Now, the general proposition states that the type of categorization system with which an individual approaches an event influences his sensitivity to that event; consequently, if the statement is valid, the sensitive person, the accurate predictor, would have approached the situation with the "proper" pattern of categories. Hypothesis 2 is concerned with the nature of these "appropriate" patterns.

Hypothesis 2

> the sensitivity of one individual for another person will increase as the degree of congruence between the categorization systems of the two individuals increases

Rationale for Hypothesis 2

An individual's categorization system is indicative of the way in which he views the world (cf. pp.12-14). If two persons tend to see the world in the same terms (with the same type of categorization system), then there can be said to exist an "empathic understanding" between these individuals.

That is, since both persons perceive events in the same form, there will be an appreciation of the thoughts and feelings, each for the other. In short, the two individuals would "see eye to eye" on things.

If one of the persons described above is the O and the other is a Judge, then it could be expected that such a J would demonstrate a high degree of accuracy in his predictions. Thus, an appropriate categorization system for a Judge in this study would be one that was similar to that of the O. As Collins (50) writes:

> ...the formation of an impression of the personality of another is a large function of the underlying perceptual-cognitive organization process in the observer.

It is highly unlikely that there can be perfect congruence between the categorization systems of two individuals; Hypothesis 2 holds that there is a direct relationship, however, between the degree of congruence of the category systems of the two individuals and the sensitivity of each for the other.

Some experimental data supporting this contention had been reported (52,76). Notcutt & Silva (76) contend:

> The greatest success in predictability should be in those areas in which the greatest similarity of behavior or attitudes exists.

The importance of the substantiation of Hypothesis 2 is suggested by Remmers (83) and Wolf & Murray (108):

> The effectiveness of clinicians and counselors in general, can, as a matter of fact, itself be measured in terms of <u>their</u> ability to empathize. (83)
>
> One might hazard a statement that without empathy a man cannot make an accurate diagnosis.... (108)

Thus, both therapeutic and diagnostic competence is seen to rest on the appropriateness of the categorization system with which the clinician approaches these situations.

To recapitulate:

General Proposition:

the type of preconceived categorization system with which one approaches events will affect sensitivity to nuances in those events

Hypothesis 1:

the sensitivity of an individual to nuances of a situation will be diminished if he approaches that situation with a preconceived categorization system

Hypothesis 2:

the sensitivity of one individual for another person will increase as the degree of congruence between the categorization systems of the two individuals increases

METHOD

Overview

In this study a group of clinical psychologists (Judges) was asked to predict the self-concept, as revealed by Q-sort methodology, of a client (Other) who was being seen in therapy. The Input upon which these predictions were based, consisted of different types of stimulus material: (i) O's therapy interviews; (ii) diagnostic protocols obtained from O. In making his predictions the Judge followed Q-sort procedures. That is, J arrayed a population of statements, indicative of personality traits, along a continuum of "most like" to "least like", as he believed the client would sort these same items. A Judge's score (Outake) was the degree of accuracy which he was able to approximate the client's actual sorting of the statements. The scores obtained by the Judges in this manner provided the dependent variable for this study. The independent variable was the sequence in which the stimulus material--interview or diagnostic--was presented to the Judges. By presenting different groups of Judges with various sequences, it was possible to test (with analysis of variance) the effect of each type of stimulus material upon the accuracy scores obtained by the Judges.

Components of the Experimental Design

The italicized words in the preceeding section represent the four major components of the experimental design employed in this investiagation (cf.26,43,45,97).[1]

The Judge --- whom the experimenter is attempting to measure

The Other --- whom the Judge is asked to investigate

The Input --- information concerning O which is made available to the Judge

The Outake -- the statements or predictions about O obtained from the Judge

The Judges

The individuals who served as Judges in this study were 32 clinical psychologists. They were employed either in the Counseling Center of a large midwestern university or by the Veterans Administration. These clinicians represent a wide range of experience ranging from trainees with limited clinical experience to staff psychologists with 18 or more years of experience.[2]

---

[1] In a further elaboration of this analysis, Gage & Cronbach (45) classify investigations in this area along the following dimensions: (a) degree of acquaintance--between J and O; (b) degree of extrapolation--amount of inference required between Input and Outake. In the present study the Js have little acquaintance with O and much extrapolation is required. In essence, the Judge must predict "how an individual deviates from the typical behavior of the particular group he belongs to--describing the unique behavior of the individual as in clinical diagnosis" (45,p.413).

[2] See discussion in Controls, p.38.

The Other

The individual, Mr. "L", who provided the stimulus material for this study was an undergraduate student at a large midwestern university. At the time that the study was undertaken he was being seen on a thrice-weekly basis by a therapist (not the author) at the university's Counseling Center. Mr. "L", upon being asked by the author, agreed to participate in this research. He was told that he would be required to take several psychological tests (administered by the author) and that recordings of several of his therapy interviews would be needed. Mr. "L" was informed that his anonymity would be preserved, and, accordingly, all identifying material was removed from the test protocols and the therapy sessions.

The Input

The stimulus materials upon which the Judges based their predictions were of two main types: (i) diagnostic protocols, and (ii) interview material.

The Diagnostic Protocols consisted of the following:

(a) the sex, age, marital status, academic class, and military status of the client (Mr. "L" was a Korean veteran)

(b) a brief statement of Mr. "L's" academic interests

(c) his percentile level on the Wechsler-Bellevue Intelligence Scale

(d) a verbatim typescript of Mr. "L's" Rorschach performance; including, scoring, summary sheet, and location chart

(e) a verbatim typescript of his performance on the Thematic Apperception Test; cards 1,2, 4,6BM,7BM,10,12,13MF,13B,17BM,18BM, and 16.

(f) the comments and Mr. "L's" associations on the Draw-A-Person test. A copy of Mr. "L's" drawings were also included in the diagnostic material given to the Judges. (The client was asked to make his drawings on Ditto master sheets so that they could be reproduced.)

The Interview Material consisted of two recorded therapy sessions--I' and I"--from which the therapist's remarks had been deleted. Verbatim typescripts, to be read while the recordings were played, were also provided for the Judges. The entire 11th and 13th therapy interviews were used.

The therapist's remarks were removed from the recordings of the interview material for the following reasons:

(a) this procedure would make the results obtained in this investigation more comparable to those reported in a similar study by Kell (59)

(b) deleting the therapist's remarks made the interview material more "pure" in that the Judge could only hear what the client said and, consequently, would not be distracted by the therapist's voice, therapeutic acumen, therapeutic orientation, etc.

(c) these deletions made listening to the interviews a more novel experience, thereby increasing the Judges' attentiveness; further, it made the session somewhat more difficult to follow which had the effect of increasing the Judges' concentration on the material

(d) this procedure reduced the time required for listening to the recordings and, therefore, decreased the amount of time needed to complete the several tasks required of the Js[4]

The method employed, and the considerations involved in selecting sessions I' and I" are described in Appendix G.

By using both recordings and verbatim typescripts of the interviews the intent was to make the stimulus material as rich as possible. Use of the one without the other would have resulted in a loss of many cues available in the total situation.

Support for the above contention is reported by Luft (68). This author finds a "distinct advantage" gained by listening to an interview as compared with reading a typescript of the same interview when Judges were asked to predict O's responses on the basis of each type of material. Tupes (104) also finds the use of both types of cue sources an advisable procedural step.

The Outake

In the sub-sections which follow the rationale for using Q-sort methodology and a description of how the Q-arrays were constructed will be given. Succeeding sub-sections will contain a description of the scores obtained by this methodology and the interpretation given to these scores.

---

[4] See discussion in Controls, p.41.

A Q-sort Methodology was utilized in this study in view of the many advantages it was felt that it possessed. As Cronbach (18) has pointed out, this forced choice procedure is more "penetrating" than the common questionaire; it eliminates response sets (6) such as saying "Yes" to all favorable items and "No" to all unfavorable ones; it is free from such idiosyncratic responses as saying "Cannot Say" twice as often as other respondants; it requires every person to put himself on the measuring scale in much the same manner; and, it frees the subject from many difficult and rather unimportant discriminations such as would be involved in ranking every statement.

Block (9) compared a forced and unforced type of sorting and reports:

> ...the forced sort is at least as stable as the unforced sort.... The forcing procedure did not apparently decrease the reliable variance present.
>
> Using a simple distance measure, the forced sort emerges as more dependable....
>
> More discriminations tend to be made in forced rather than unforced Q-sorting situations. (9,p.484)

Q-sort methodology also possesses a computational advantage. Since Q-arrays have identical means and standard deviations (a consequence of the forced distribution), it is possible to compute Personian product-moment rs by the quick, efficient method of difference[5] procedure.

---

[5] (16) $$r = 1 - \frac{D^2}{2N\sigma^2}$$

Of the 32 Judges, only eight had previously participated in research involving Q-sorting. Thus, this methodology had the additional advantage of novelty which served to enhance the interest displayed by the Judges.

The Q-arrays[6] used in this study--Q' and Q"--each contained 70 statements. These items were derived from Edwards' (30) Personal Preference Schedule (PPS), which, in turn, was based on Murray's (75) system of needs. The decision to develop this particular item-population was based on the following considerations:

Edwards has demonstrated (29,31) a .87 correlation between the readiness to subscribe to an item as descriptive of one's self and the social desirability of that item. In view of this finding it was deemed inadvisable to ignore social desirability as a factor capable of influencing self-descriptive data. Consequently, an attempt was made to equate the two Q-arrays for social desirability.[7]

Since an equating on this variable had been achieved in the PPS, these items were selected. (Edwards' statements

---

[6] See discussion in Controls, p.41.

[7] The writer attempted to obtain Edwards' original data in order to carry out this equating, but was unable to do so. In a personal communication Dr. Edwards referred the writer to a study by C. James Kleet (63). This work was obtained but, unfortunately, it did not contain the social desirability values for all the items. Consequently, the two Q-arrays were equated by a less precise method (see Appendix B).

were modified slightly--usually by eliminating the words "like to"; e.g., "I (like to) participate in new fads and fashions"--in an effort to make these items more discriminable for the sorters.)

A second consideration, which seemed to make these items particularly suitable, was the fact that they were based on Murray's personality theory. In view of the nature of this study--particularly its diagnostic aspects--it was advisable that items be used which were readily translatable into the types of hypotheses arrived at from diagnostic protocols. Moreover, the inclusion in the test battery of the TAT--which was specifically designed to elicit expression of these needs--seemed to make these particular items highly applicable.[8]

The two Q-sorts used in this study each contain 70 items. Sixty of these items were derived from the PPS. That is, each of Murray's 15 need categories is represented by four items in each Q-array. Ten additional items were added to each array since, in reviewing Stephenson's book on Q-methodology (95), it was noted that rarely was a statement population of less than 70 items utilized. Further, in selecting these additional items it would be possible to partially obviate a frequent criticism made of personality

---

[8] It is possible that these items introduce a bias in the direction of aiding predictions based on the daignostic material; see discussion in Limitations, p.64.

tests of the nature of the PPS. The question is frequently raised regarding the validity of the statements employed--do they actually represent the needs to which they are assigned.

With these considerations in mind, items were adopted from Fordyce's dissertation (37). These statements had been selected for his study on an empirical basis[9], as being representative of n (need) Independence and n Dependence; they also had been scaled for social desirability. By incorporating these items into the Q-arrays the social desirability balance was maintained and the sorts enlarged.

In summary, then, each of the Q-sorts contained 70 items. In each array, for each of Murray's need categories, there were four items. Two additional need categories--Independence and Dependence--were represented by five items each, in both arrays. Thus, two Q-sorts, balanced for social desirability and possessing the advantageous "lumpy" quality (18)--clusters of items--were developed; Appendix D lists the items of both Q-arrays.

The Scores[10] obtained using Q-methodology depend on the

---

9 Fordyce selected his statements in the following manner: 307 items were submitted to seven clinical psychologists who classified the statements as to "independency" and "dependency" (the clinicians supplied their own definitions of these terms). Two hundred and twenty-five (225) of the items were classified identically by five of the seven Judges. These items were then rated for social desirability by 68 college students.

10 The data of this study consist of correlation coefficients which were transformed into "scores" by using Fisher's z' transformation; cf.28,p.126-8).

number and kinds of arrays the Judges are required to make.

In this investigation each Judge did three sortings:

i. a self-sort--in which the J distributed the items along a continuum of "least like" to "most like" for himself

ii. a predictive-sort--in which J placed the statements along a "least like" to "most like" continuum as he believed O would have sorted these same items

iii. a second predictive-sort--as in (ii) with J using the alternative from Q-array[11]

From these sortings several scores were attainable:

a. Accuracy (ACC)--also called "prediction score" (34) and "raw empathy" (7,53)--this is the correlation between J's predictive sort and O's actual sort

b. Real Similarity (RS)--cf.34--the correlation between J's self-sort and O's self-sort

c. Attributive Similarity (AS)--also known as "projection" (7,53); "assumed similarity" (42); and, "non-differentiation" (4)--correlation between J's self-sort and his predictive-sort

For example, Judge A (see p.37) would have the following scores; the notation ":" means "correlated with".

a. ACC--(1) predictive Q' : O's Q'
(2) predictive Q" : O's Q"

b. RS---J's self-sort (Q") : O's Q"

c. AS---J's self-sort (Q") : J's predictive-sort (Q")[12]

---

[11] Each Judge used both Q-arrays in one of two sequences: (i) Q", Q', Q"; or, (ii) Q', Q", Q'; cf. discussion in Controls, p.41.

[12] A second AS score could be obtained by correlating J's self-sort (Q") with J's predictive-sort (Q') using the 17 "needs" as the data to be correlated; cf. discussion in Results and Discussion, p.45.

The Interpretation of the Scores was as follows:

ACC is essentially "understanding the other person in the diagnostic sense of the word" (34,36). Gage, et al (46) list the following terms for ACC: diagnostic ability, empathy, interpersonal perception, sensitivity, social perception, and understanding. These terms have a common element in that they convey the idea of "clinical intuition" to use Hathaway's expression (54). In this study the term "accuracy" will be used as indicative of sensitivity. That is, the higher the ACC the greater the sensitivity displayed by the Judge.

RS is defined by Fiedler (34) in the following manner:

> "Real" in this case indicates, of course, only the real similarity between self-descriptions, and only to the extent to which the self descriptions reflect personality structure can we speak of real similarity between the persons involved.

AS has been used frequently in studies of this type, but, as the plethora of other names may suggest, it is the most difficult of the three scores to interpret. A more extensive discussion of this score and the reasons for its exclusion from this study are given in Appendix H.

Procedures

Since the procedures for O and J differ, each will be presented in a separate sub-section; a summary of these procedures, in tabular form, will conclude this section.

Procedure for Other

The recorded interviews were obtained in the manner described in Appendix G. In the testing session O was given a sheet containing suggested instructions for doing a Q-sort; Appendix C. He did a self-sort using the items in Q'; he then did a second self-sort using the items in the Q" array. The required distribution for these sortings is shown in Table I .

TABLE I

REQUIRED FREQUENCY DISTRIBUTION
FOR ALL Q-SORTS

Pile Number	1 LEAST LIKE	2	3	4	5	6	7	8	9 MOST LIKE
Number of Items	1	3	7	15	18	15	7	3	1

Following the second self-sort O was administered the projective tests (Rorschach, TAT, DAP). O's final task was to rate himself on each of the 17 need categories on a nine point Rating Scale [13]; Appendix E.

Procedure for Judges

Each Judge was asked to complete an Experience Data Sheet (Appendix F). On the basis of the information obtained

[13] This scale was composed of the definition and list of defining adjectives used by Murray (75) to characterize each of his need categories.

he was assigned[14] to one of the experimental conditions; this determined the sequence in which the stimulus material--diagnostic or interview--was presented to the J (see below).

After completing the Experience Data Sheet, each Judge was given the suggested instructions for doing Q-sorts (Appendix C). The J then did a self-sort. Following the presentation of the first stimulus material, the Judge did his first predictive-sort. Then the second stimulus was presented and J did a second predictive-sort.

Following this third sort, the Judge was given the Rating Scale (Appendix F) and was asked to make a "diagnostic" rating of Mr. "L". That is, the Judges were instructed: "Now, on this sheet, rate the subject as you think he really is."

To summarize, the procedures employed were:

Other
- i. self-sort with Q' items
- ii. self-sort with Q" items
- iii. projective tests
- iv. self-rating on Rating Scale

Judge
- i. complete Experience Data Sheet
- ii. self-sort*
- iii. first stimulus material presented
- iv. first predictive-sort*
- v. second stimulus material presented
- vi. second predictive-sort*
- vii. diagnostic rating of 0 on Rating Scale

* Sequence (i) or (ii), page 33, was followed

---

[14] Originally the intention was to match Js on all items listed on the Experience Data Sheet. However, this procedure was infeasible due to the limited number of clinicians available to the writer. Matching was based on only therapeutic and diagnostic experience; see discussion in Controls, p. 38.

Experimental Conditions

The four experimental conditions employed consisted of different sequences for the presentation of the stimulus material. These sequences were:

i. Interview material followed by Diagnostic protocols (I"D)

ii. Diagnostic protocols followed by Interview material (D I")

iii. the 11th Interview followed by the 13th Interview (I'I")

iv. the 13th Interview twice (I"I")

Figure 1 presents these sequences in schematic form. In addition to the symbols already introduced, I stands for "predictive Q-sort based on Interview material"; D means "predictive Q-sort based on Diagnostic protocols"; capital letters represent the Judges. Session "two" immediately followed session "one".

		ONE		Sessions		TWO	
Experimental Conditions		Q'	Q"	'		Q"	Q'
	i.	A--D	E--H	I"	D	A--D	E--H
		Q'	Q"	'		Q"	Q'
	ii.	I--L	M--P	D	I"	I--L	M--P
		Q'	Q"	'		Q"	Q'
	iii.	R--U	V--Y	I'	I"	R--U	V--Y
		Q'	Q"	'		Q"	Q'
	iv.	AA--DD	EE--HH	I"	I"	AA--DD	EE--HH

Figure 1. Schematic presentation of procedures for all Judges; N = 32, n = 8.

Experimental conditions I"D and D I" (i and ii in the above schema) indicate that the interview and diagnostic material were presented to the Judges in a counter-balanced sequence. This design permitted the evaluating of the effects of one type of stimulus material on the other.

Experimental condition iii--I'I"--presented the Judges with the same kind of material (all interview). This was an attempt to control for the "heterogeneity" of the stimulus material in conditions i and ii, while keeping the "amount of information" given the Judges roughly the same.

Experimental condition iv--I"I"--presented the Judges with the same material on two occasions. This was an attempt to control for the fact that the Judges in the other conditions were given different material to base their predictions on, while keeping the "amount of exposure" to the stimulus material roughly the same.

Additional Control Features

In addition to the controls incorporated in the experimental conditions, an attempt was made to control other factors in an effort to increase the rigor of the investigation. These factors, their rationale, and the procedures utilized will be detailed in this section.

(a) Experience--Kell (59), in a study somewhat similar to this one, used experienced therapists, pre-practicum students, and chemists (n=16). He found that there was a

significant difference between the mean accuracy of the predictions made by these groups. In "step-wise" fashion, the therapists were significantly more accurate than the students, who, in turn, were significantly better than the chemists.

In view of this finding, and the wide range in therapeutic (and diagnostic) experience present in the population employed in this study, it was deemed necessary to control the experience factor. Since diagnostic protocols as well as therapy interviews served as stimulus material in this study, an attempt was made to take cognizance of both types of clinical experience.[15]

TABLE II

RANGE OF JUDGES' CLINICAL EXPERIENCE UNDER FOUR EXPERIMENTAL CONDITIONS

Exp. Condition*	n	Therapy**	Diagnosis***
I"D	8	37 - 5000	12 --- 1100
D I"	8	0 - 5000	6 --- 2000
I'I"	8	0 - 4500	30 --- 800
I"I"	8	35 - 5000	0 --- 1000

* See Experimental Conditions, p.37ff.
** No. Individual hours conducted and supervised
*** No. Diagnostic reports completed and supervised

[15] Based on information Js provided on Experience Data Sheet, each was assigned to one of the experimental conditions; care being taken to have same number of highly experienced people in each group.

Since the data in Table II is of an approximate nature, in order to ascertain whether the groups of Judges differed significantly on the therapy- and diagnostic-experience variables, a non-parametric test was employed--Kruskal-Wallis H-test (analysis of variance using ranked data) (91). The results of this analysis[16] led to the conclusion that the four groups did not differ significantly; i.e., the groups of Judges were reasonably well equated in terms of amount of therapy- and diagnostic-experience.

(b) Time Spent with Stimulus Material--Since two kinds of stimulus material were used--diagnostic and interview--in order that differences in ACC which might occur could be properly attributed to the stimulus difference, it was deemed necessary to control this "exposure" factor. That is, if the Judges in one experimental condition spent less time with, say, the interview material, then it would be unclear whether differences which might occur between the ACC scores of this group and the scores of another group, were due to the exposure factor or the experimental condition.

In accord with the above considerations, a minimum exposure time was set by the writer. The playing of the deleted-therapy-interviews (cf.p.27ff.) required approxi-

---

[16] The obtained values for the therapy- and diagnostic-experience variables were -.956 and -.3.0397, respectively. For three degrees of freedom a value of 7.815 is required for significance at the .05 level (11.341 at .01).

mately 20 minutes. A Judge was not permitted to begin sorting until after the interview had been played through. This same minimun time--20 minutes--was used with the diagnostic material. Each J was told he might spend 20 minutes, even longer if he wished, before beginning his predictive-sort. It was possible to maintain this minimum exposure time by withholding the pile of Q-statements until this time had elapsed.

(c) Fatigue--The modal time needed for the Judges to complete all the required procedures was two hours (the range was from 85 minutes to three hours and 20 minutes). While little could be done to control these individual differences, it was felt that procedural steps which would help decrease this amount of time might keep fatigue at a minimal level.

The decrease in the amount of time required for hearing the therapy sessions engendered by deleting the therapist's remarks from the recordings, was one method used to control fatigue. It was felt, also, that these deletions increased the novelty of the procedures and that this too would serve to keep fatigue at a lower level.

Another innovation, which also had the goal of helping to reduce fatigue, was the use of two equated, though different Q-arrays. It was hoped that alternating the packs of Q-statements would decrease boredom and, thereby, fatigue.

(d) Memory--The use of two Q-arrays, an experimental device unique to this investigation, also served a function

other than that mentioned above. Previous studies are reported in the literature in which the Judges or Others were required to do more than one Q-sort--e.g., the self-, ideal self-, and ordinary person- sorts that are frequently used. However, in these investigations the same population of statements is utilized for all sortings. This procedure introduces an uncontrolled, possibly differentially operative, confounding factor; namely, memory.

In such an experimental design, a Judge, possibly remembering some of the item placements of his first sorting, might consciously attempt to array these same items either the same or differently in subsequent sorts. It becomes unclear whether the correlations which would be obtained under these circumstances are due to the experimental conditions employed, or are a function of the Judge's memory --be it accurate or not--or some combination of the two. The literature does not contain any introspective data obtained from Judges on this issue.

In introducing a second, equivalent, population of Q-statements, the intent is to:

1) minimize, to some degree, a confounding memory factor

2) maintain interest at a higher level by deminishing the effects of boredom

3) permit the obtaining of a reliability measure for O

It was these considerations which led to the decision to develop two Q-arrays. The two piles of statements were given to the Judges in a mixed sequence--(i) Q", Q', Q"; or, (ii) Q', Q", Q'--with the expectation that this procedure might interrupt any memorization process.

In an effort to further minimize the differential influence of memory, a second control was employed. The Js were permitted to take notes during the time they were exposed to each type of stimulus. They retained these notes, as well as the printed stimulus material, while doing their predictive Q-sorts. In other words, a Judge was free to "refresh" his memory of the stimulus material while making his predictions.

These same conditions prevailed for both prediction sessions; however, following session one, all notes and printed materials were removed prior to the presentation of the stimulus material for prediction session two.

(e) Interaction between Q-sorts--Using two Q-arrays emphasizes the possibility that the items in the first array may interact with those in the second one. This possibility is also present when the same population of Q-statements is used for all the sortings, but it appears to have been relatively ignored as a source of "error" in previous researches using this methodology.

Through the use of a counter-balancing technique, an

attempt was made to "cancel out" interaction effects. That is, four Judges in each experimental condition used sequence i (see section "d" above) and the other four Judges used sequence ii. Since the main concern of this investigation was with group differences produced by the different experimental conditions, this procedure served to control possible interaction effects by permitting such effects to operate to the same degree in each of the experimental conditions.

(f) Social Desirability of Items in Trait Population--

See discussion on page 30ff. and Appendix B.

RESULTS AND DISCUSSION

Cronbach (18) points out that data obtained with Q-arrays in which there are "clusters" of items are susceptible to two different types of analysis; correlations may be computed on the individual items, or the "scores" for each cluster can serve as the basis for the correlations. The latter method gives more reliable measures and is preferred.

In the present study the clusters of statements represent need categories. The analysis was conducted on the correlation of these "need scores".[1] That is, the mean of the item placements for a particular need category was taken as the score for that need.[2] Thus, O and each Judge had 17 scores for each Q-sort. Table IV (p.49) shows the results of the computations carried out on these cluster scores.

Reliability

Other

A reliability coefficient for O was computed using an analysis of variance technique. Such a procedure makes it

---

[1] Correlations were also computed on an item basis. The results of this statistical treatment are given in Appendix J.

[2] This procedure, although routinely employed in studies of this nature, makes a tacit assumption regarding the clusters of items; namely, that each cluster comprises a sub-test in the total population of items and, further, that the integrity of these sub-tests is not destroyed when the cards (on which the items are printed) are shuffled prior to being sorted.

possible to ascertain whether or not the test measures with sufficient accuracy to differentiate between the needs; and, also whether or not the practice effect is significant. Table IIIshows the results of these computations.

TABLE III

RELIABILITY FOR "OTHER"

SOURCE OF VARIANCE	SUM OF SQUARES	df	MEAN SQUARE	F	p
Practice Effects (Between Q' & Q'')	.0002	1	.0002	---	
Between Needs	17.4644	16	1.0915	2.6364	.05>p>.01
Error	6.6251	16	.4140		
TOTAL	24.0897	33			

$$r_{tt} = \frac{\text{Between Needs} - \text{Error}}{\text{Between Needs} + \text{Error}} = .450$$

Admittedly, the reliability coefficient obtained for O is low; however, the analysis has demonstrated that the Q-sorts, as used by the O, differentiate between the various need categories at better than .05 level. Regarding the advisability of using a criterion with low reliability, Thorndike writes:

> ...it is not of <u>critical</u> importance that the reliability of a criterion be <u>high</u> as long as it is established as definitely greater than

> zero. Even when the reliability of a criterion is quite low, given that it is definitely greater than zero, it is still possible to obtain fairly substantial correlations between that criterion and reliable tests and to carry out useful statistical analyses in connection with the prediction of that criterion. (101,p.106)

Since an r = .450 does not reach the value (.468) required to be significantly different from zero at .05 level (N = 17), the results of this study become more equivocal.[3]

Judges

The consistency with which the Judges used the two Q-arrays was computed following a technique outlined by Ebel (27). The coefficients obtained--.737 and .796 for Q' and Q", respectively--are each based on the first predictive sorts of eight Judges; all 16 Judges made their predictions on the basis of I".

The 16 Js used in computing these reliability coefficients were the eight assigned to the I"D condition and the eight assigned to the I"I" condition. The first four Js in each of these groups made their initial predictions using Q'; these sorts were used to compute the reliability of Q'. The second four Judges in each of these groups made their first predictive sorts using Q"; accordingly, the reliability for Q" was based on these arrays.

---

[3] A reliability coefficient for O based on the individual items was also computed; r = .197. Thus, the advisability of using the "need scores" and the validity of Cronbach's recommendation (p.45) was demonstrated.

The reliabilities obtained in this study are lower than those found by Frank (39). This author, using a test-retest methodology, reports correlations ranging between .93 and .97 for self-sort with Q-statements. The lower reliabilities found in this investigation may be the result of (i) basing reliabilities on predictive-sorts rather than self-sorts; (ii) using two trait-populations rather than the same one for all sortings.[4]

Results

In order to properly evaluate the data obtained in this investigation it is necessary, first, to ascertain whether or not the Judges were able to predict O's self-sort with a degree of accuracy greater than chance. If the Judges were unable to exceed a chance level, then any further statistical analysis would be open to serious questioning.

The correlations reported in Table IV (p.49) reveal that of the 64 ACC scores, 20 were significantly different from zero at the .05 level (12 of these at .01).

A further analysis of this data, shown in Table V (p.50), indicates that 16 Judges had at least one ACC score significantly different from zero at the .05 level; 11 Judges were able to reach the .01 level on at least one prediction.

---

[4] See discussion in Controls, p.41ff.

TABLE IV

ACC AND RS CORRELATIONS
FOR ALL JUDGES

GROUP	JUDGE	SESSION ONE ACC		SESSION TWO ACC		GP. MEAN ACC	RS
	A	.032		-.113			.563**
	B	.717**		-.029			.530*
	C	.412		.447			.502*
	D	.557*		.352			.393
I"D	E	-.042		-.058			.523*
	F	.783**		.465			.545**
	G	.601**		.331			.375
	H	.108		.023			.208
	Mean	.444**	---	.118		.312**	
	I	-.274		.080			.472*
	J	.082		.121			.557*
	K	.347		-.024			.272
	L	.593**		.388			.643**
D I"	M	.759**		.104			.295
	N	.658**		.535*			.406
	O	.449		.456			.691**
	P	.645**		.475*			.733**
	Mean	.450**	---	.279**		.367**	
	R	.639**		.299			.533*
	S	.567*		.536*			.735**
	T	.581*		.174			.279
	U	.094		.158			.750**
I'I"	V	-.217		.433			.183
	W	.667**		.376			-.005
	X	.733**		.628**			.380
	Y	.543*		-.033			.034
	Mean	.482**	---	.337**		.416**	
	AA	.641**		-.081			.256
	BB	.407		.415			.609**
	CC	.079		.429			.114
	DD	.105		-.045			.135
I"I"	EE	.559*		.205			.669**
	FF	-.106		.429			.655**
	GG	.298		.368			.293
	HH	.391		.122			.036
	Mean	.316**	---	.239**		.278**	
TOTAL MEANS		.427**		.262**		.347**	

* Significantly different from zero at .05 level
** Significantly different from zero at .01 level

TABLE V

NUMBER OF JUDGES ACHIEVING ACC SCORES DIFFERENT FROM ZERO AT .05 AND .01 LEVELS OF SIGNIFICANCE

Level of Significance	Number of ACC Scores at this Level		
	0	1	2
.05	16	12	4
.01	21	10	1

It is difficult to evaluate the frequencies reported in Table V since there is no way of determining how many ACC scores, different from zero at a specified confidence level, could be expected from the group of Judges by chance alone. Moreover, the question arises as to the magnitude of ACC scores which could be expected to occur by chance. In an attempt to establish a base level from which the obtained ACC scores might be evaluated, a "robot" was constructed.

"Robot" Accuracy

A self-sort and two predictive sorts, using the two Q-arrays in sequence i--Q", Q', Q"--were made for the Robot. That is, for each sort the pile of cards containing the 70 Q-statements was shuffled and arrayed into the required frequency distribution (cf.p.35) according to the dictates of a table of random numbers. The obtained scores were:

ACC (Q') .017 RS -.012

ACC (Q") -.103

Following the procedures outlined by Edwards (28,pp.128-131) fiducial limits at the .05 level were established for the Robot's scores. These were:

ACC	(Q')	+.356	to	-.338
ACC	(Q")	+.276	to	-.432
RS	(Q")	+.342	to	-.360

If the ACC scores in Table IV were evaluated with the fiducial limits of the Robot's scores in mind, the results are as shown in Table VI.

TABLE VI

NUMBER OF JUDGES ACHIEVING ACC SCORES BEYOND ROBOT'S FIDUCIAL LIMITS*

Number of Times Robot's Scores Exceeded	Number of Judges
0	8
1	11
2	13

* Fiducial Limits established at .05 level

The results shown in Table VI demonstrate that 75% (24 out of 32) of the clinicians, who served as Judges for this study, were able to predict the self-sort of a client, at least once in two trials, significantly better than chance; slightly better than 40% of the Js exceeded chance on both predictive sorts.

In addition, there was a highly significant degree of agreement among the Judges when they rated C "diagnostically" (cf.p.35ff.). These ratings were ranked and Kendall's (91) coefficient of concordance computed. The value obtained, 147.24, for 16 degrees of freedom, is significant beyond the .001 level.

A corrected rank-difference correlation of .419 was found between C's own ranked ratings on the Rating Scale and the consensus ranking of the 32 Judges. This value does not reach the level required for a significant difference between a rank correlation and zero (102). The magnitude of this correlation, however, suggests that the role played by diagnostic ability in a Judge's ACC score would be a fruitful area for investigation.

The above results were interpreted as indicating that the Judges in this study predicted better than the level which could be expected by chance alone. Further, there was a significant degree of agreement among the Judges in the impression they gained of O from the divers stimulus material.

Hypothesis 1

Hypothesis 1 holds that an individual's sensitivity to a situation is diminished when he approaches that situation with a preconceived categorization system. In attempting to demonstrate the validity of this statement it is necessary

to take into account three variables involved in the study:

i. Experimental Session -- 'one' and 'two'

ii. Type of Material -- diagnostic and interview

iii. Sequence of Material -- heterogeneous (I"D ;D I")
homogeneous (I'I")
identical (I"I")

Support for Hypothesis 1 would obtain if the ACC scores in session 'two' were significantly lower than those in session 'one'. Since it was possible for interactions to occur between the three variables, analysis of variance was selected as the appropriate statistical technique. Through this type of analysis it was possible to test all the variables and their interactions, simultaneously. Following the method outlined by Edwards (28,pp.288-296) for analysis of variance with repeated measurements on several independent groups, the results shown in Table VII (p.54) were obtained.

The F between sessions of 7.3544, for 1 and 28 degrees of freedom, is significant between the .05 and .01 levels. An examination of Table IV (p.49) reveals the total mean for session 'one' to be higher than the mean of session 'two'. Consequently, it was concluded that the categorization system which the Judges brought into the second prediction session led to significantly lower ACC scores.[5] To the extent that

[5] The ACC scores of 23 of the Js in this study showed this trend. Of the nine Judges whose predictions were in the opposite direction (i.e., the second prediction more accurate than the first) not one of their ACC scores was significantly different from zero.

these ACC scores reflect the Judges' sensitivity, Hypothesis 1 has been demonstrated.

TABLE VII

ANALYSIS OF VARIANCE OF ACC SCORES FOR FOUR GROUPS OF JUDGES UNDER DIFFERENT EXPERIMENTAL CONDITIONS WITH TWO SESSIONS FOR EACH GROUP

SOURCE OF VARIATION	SUM OF SQUARES	df	MEAN SQUARE	F
Between Conditions	.2209	3	.0736	---
Between Js in same Condition	3.6083	28	.1288	
Total Between Js	3.8292	31		
Between Sessions	.5685	1	.5685	7.3544
Interaction: Sessions x Conditions	.0831	3	.0277	---
Interaction: Pooled Js x Sessions	2.1652	28	.0773	
Total Within Js	2.8868	32		
TOTAL	6.6460	63		

The finding of a significant difference between sessions, although in accord with the gating principle in Bruner's (11) general formulation, is also embarrassing to the part of his theory dealing with the "closer look". Bruner contends that permitting an individual to take a "closer look" at a situation is one of the antidotes for non-veridical perception. Experimental Condition I"I", in essence, provided eight of

the Judges with this "closer look". As Table IV (p.49) reveals, the group mean for I"I" is higher for session 'one' than it is for session 'two'. While this difference in not statistically significant (t = 1.3735; df = 7; .50>p>.20) it is in the opposite direction to Bruner's theorizing. According to the "closer look"principle, the Judges should have higher ACC scores in session 'two'.

The data of this study appears to be more in agreement with Hathaway's statement:

> It would seem that the consistant error of both individual participants and of groups somehow derives from having too much information--at least too much information desturbs the predictive process. ...Of course the arguement would be absurd if the additional information were directly informative about the items to be predicted. (54)

In an attempt to isolate the experimental conditions contributing to the obtained significant difference between the two predictive sessions, separate analyses of the individual experimental conditions were conducted. At the same time, the possible effect of different levels of therapeutic and diagnostic experience was investigated; these tables are given in Appendix I. Parenthetically, it may be noted that the significant relationship between level of experience and predictive ability reported by Kell (59) was not replicated in the results of this analysis.

The tables in Appendix I reveal that no one experimental

condition resulted in a statistically significant difference between session 'one' and session 'two'. However, the Fs obtained for the I"D and D I" conditions are consistantly higher than those found in the I'I" and I"I" conditions. In an attempt to ascertain whether the heterogeneous or the homogeneous material differentially affected the Judges' performance in the two sessions, two statistical analyses were conducted--Tables VIII and IX.

TABLE VIII

ANALYSIS OF VARIANCE OF ACC SCORES
APPLIED TO I'I" AND I"I" CONDITIONS

SOURCE OF VARIATION	SUM OF SQUARES	df	MEAN SQUARE	F
Between Conditions (I'I" & I"I")	.1985	1	.1985	2.4320
Between Js in same Condition	1.1446	14	.0817	
Total Between Js	1.3431	15		
Between Sessions	.1448	1	.1448	1.5210
Interaction: Sessions x Conditions	.0204	1	.0204	---
Interaction: Pooled Js x Sessions	1.3328	14	.0952	
Total Within Js	1.4980	16		
TOTAL	2.8411	31		

Neither of the Fs obtained in Table VIII, for 1 and 14 degrees of freedom, reach the value (4.60) required for significance at the .05 level.

Thus it would appear (and Table IX bears this out) that the lowered predictive accuracy in session 'two' is a consequence of the interference produced by the presentation of heterogeneous stimulus material. In order to investigate more closely the divers effects of the remaining two variables (p.53) on the heterogeneous conditions, a statistical methodology for counter-balanced tests (94) was employed. Table IX shows the results of this analysis.

TABLE IX

ANALYSIS OF VARIANCE FOR COUNTER-BALANCED TESTS APPLIED TO I"D AND D I" CONDITIONS

SOURCE OF VARIATION	SUM OF SQUARES		df		MEAN SQUARE	F
Independent Results:						
Presentation order	.0221		1		.0221	---
Residual Between Js	2.4637		14		.1759	
Total Between Js		2.4858		15		
Correlated Results:						
Between I" and D	.0160		1		.0160	---
Between Sessions	.4703		1		.4703	7.9175
Residual Within Js	.8325		14		.0594	
Total Within Js		1.3188		16		
TOTAL		3.8046		31		

As shown in Table IX, accuracy of prediction is not significantly affected by the sequence in which the material is presented,[6] i.e., whether I" preceeds or follows D; nor is it affected by the type of material upon which the prediction is based--I" or D. The F for Between Sessions of 7.9175, for 1 and 14 degrees of freedom, is significantly different from zero at the .05 level. Since this result is based on the effects of making two predictions on each Judge, it becomes apparent that it is something within the J which accounts for the obtained result. To the extent that this interference reflects the operation of a gating process, it may be concluded that the phenomenon is most pronounced when there is a transition from interview to diagnostic material (or vice versa). Thus, this statistical procedure lends added weight to the contention that the differences found between prediction sessions are the consequences of a gating process.

---

[6] An alternative statistical analysis, suggested by Dunnett's (23) procedure for comparing several treatments with a standard, provided additional support for this conclusion. The mean ACC for each session in each of the experimental conditions was computed. In three of the conditions the ACC based on I" is preceeded by D, I', and I"; in the remaining condition (I"D) the ACC based on I" was the first prediction made. Thus it becomes possible to compare the effects of three different types of preceeding material on I" with a standard--I" as the first prediction.

The largest difference between a treatment mean and the standard was .2335 (for I"I"). Since a difference of 3.1390 is required for a one-tailed test of significance at the .05 level, it was concluded that none of the three sequences had a differentially significant effect on ACC.

Hypothesis 2

Hypothesis 2 states, in essence, that if two people "see" themselves in a similar way, then the amount of sensitivity of these individuals, one for the other, will be greater than if this initial similarity did not exist. RS was interpreted as indicative of the amount of similarity existing between O and J (cf.p.34). Thus, the confirmation of Hypothesis 2 would require that a significant correlation be found between ACC and RS; i.e., the more sensitive the J the higher his RS score.

Since the analysis of the ACC data revealed that there was a significant difference between the scores in the two prediction sessions, a mean ACC score was computed for each Judge; these means were then ranked. A rho correlation between these ranks and the ranks based on the Judges' RS scores was calculated. The obtained value, $\underline{r}$ = .201, did not reach the required value for a rank correlation (N=32) to be significantly different from zero (102).

In an attempt to ascertain whether or not RS was significantly related to the extreme ACC scores, a biserial correlation was computed. On the basis of their mean ACC scores the Judges were separated into two groups--the upper and lower 27% (n=9). A biserial $\underline{r}$ was computed from the RS scores of these 18 Judges. The obtained value, $\underline{r}$ = .207, was not significantly different from zero.

Although both of these results are in the predicted direction, in view of the small magnitude of the obtained correlations, it must be concluded that RS is not a significant factor contributing to high ACC scores. Thus, the results of this analysis do not lend statistically significant evidence for the validity of Hypothesis 2.[7]

To recapitulate, the statistical analysis employed in this study has led to the following conclusions:

> Diagnostic material does not markedly enhance nor significantly decrease the sensitivity of a Judge to Interview material; and, similarly, for the effect of Interview material on Diagnostic material
>
> Although the presentation sequence of the stimulus material does not significantly effect sensitivity, making two predictions, particularly if based on widely differing types of material, results in significantly less accuracy for the second prediction
>
> Support was found for the validity of Hypothesis 1
>
> The validity of Hypothesis 2 remains in question; the results were in the anticipated direction but did not reach the level of statistical significance
>
> The general proposition underlying this study, was, in the main, supported by the above findings

---

[7] The experimental procedures employed in this study may not have tested this hypothesis adequately, since the possible influence of threat was not controlled. That is, it may may have been the case that some of the obtained inaccuracy in ACC is attributable to a disinclination on the part of the Judges to "see" (predict) C as similar to themselves; cf. discussion p. 68.

Limitations

As stated previously, a broad purpose of this investigation was to examine certain variables that might be related to a therapist perceiving his client more (or less) accurately, and which, in turn, would influence the effectiveness of the therapeutic relationship. We have not, however, studied psychotherapy directly. Consequently, the results of this investigation are applicable to psychotherapy only to the extent that certain assumptions are accepted.

A first assumption is that an awareness on the part of the therapist of his client's phenomenological frame of reference is an important variable in effective therapy.

Such an assumption is directly in accord with the tenets of Client-Centered therapy, as outlined by Rogers (84,85), in which great emphasis is placed on the necessity of an empathic understanding existing between therapist and client.

From a more analytical orientation, Frieda Fromm-Reichmann writes:

> We know that the success or failure of psychoanalytic psychotherapy is...greatly dependent upon the question of whether or not there is an empathic quality between the psychiatrist and the patient. (41,p.62)

Otto Fenichel, writing from a classical psychoanalytic viewpoint, cites as a first rule regarding interpretation: One should always start the interpretation at the surface" (32,p.44). He goes on to say:

> We interpret, as is well known, what is already in the preconscious--and just a little bit more--which thereby becomes capable of entering consciousness. (32,p.53)

Whether it serves as a basis for "empathic understanding" or for determining what constitutes "just a little bit more", the foregoing suggests that the phenomenological world of the client is regarded as an important variable by major theoriticians writing on psychotherapy.

A second assumption, concerning Q-methodology, is required before the results of this study can be interpreted as having meaning regarding effective psychotherapy. It was assumed in this research that Q-sorting afforded a means for obtaining a quantifiable, flexible (18), highly complex and multi-dimensional (64) view of a client's phenomenological world. Several studies in this area (38,64,87) find that Q-sorts can adequately reporduce the phenomenology of the sorter; and, further, that a Q-sort does seem to be able to say what the sorter wants it to say despite his doubts that his impressions are "coming through".

To the extent that one is willing to grant the validity of the above assumptions, the results cited in the preceeding section have relevance regarding the usefulness of pre-therapy formulations (based on diagnostic tests)in enhancing therapeutic effectiveness. Nevertheless, these results must be interpreted in light of the limitations imposed by the experimental design employed in this investigation.

In this study many Judges attempted to predict the self-concept of one O. The procedure restricts the degree to which these results may be generalized. Questions might be raised regarding the "typicality" of O; i.e., was the O used in this study typical of the clients seen in therapy? Would the same results obtain if a female client had been used? If female Judges had been used? What if a more disturbed, or a more reliable client had been selected; would the same conclusions have been reached? Suppose the Judges were required to predict the self-sorts of two very "different" Os; would this have changed the results? The present study provides no answers for these questions.

The moderate reliability of the Q-arrays used in this investigation imposes another limitation on the obtained results. One wonders whether the magnitude of the ACC scores might not have been higher with more reliable instruments. If so, could this have changed some of the findings? Might not some of the obtained insignificant Fs reach the level of statistical significance with a more reliable measuring instrument? In other words, the relative unreliability of the Q-sorts may have produced Type II errors (cf.28,p.29-30).

A further limitation in this study involves the Judges. It has been pointed out (46) that correcting scores through the use of "intermediary keys"--e.g., J's stereotype of the "typical" client--can make studies of this nature more rig-

orous. Cronbach & Gleser (19) suggest statistical procedures which can be used when Js are asked to predict the self-sort of a "typical client", along with their other predictions.

This procedure was not employed in this study. It was deemed inadvisable to require the Judges to make still another Q-sort; the time factor, which tended to increase the likelihood of disruptive effects due to fatigue, was becoming too formidable. Moreover, the issues raised by Baker & Sarbin (4) regarding the rationale for "adjusting" ACC scores for AS (see Appendix H) seemed equally applicable regarding the use of any "intermediary key".

Still another possible source of bias concerned the use of Q-statements based on Murray's system of needs. Since these items may have been more readily translatable into "diagnostic hypotheses", this could have made prediction from the interview sessions more difficult.

It would have been possible to use items taken from therapy interviews for the Q-statements--e.g., the trait population devised by Butler & Haigh (85)--but this may have biased the results in the other direction. Obviously, Q-items derived from both sources would have resolved this difficulty, but such a population of statements was not available to the writer.

The decision to use the statements which were employed in this study was based on the following considerations: (a) the statements were available to the writer; (b) the statements

had been scaled for social desirability; (c) bias in favor of the "diagnostic prediction" was seen as less damaging.

The rationale for reason (c) was as follows: The Judges were asked to predict the self-sort of a client. If one thinks in terms of "levels of consciousness", the material contained in the interviews is "closer" to the level at which O would operate in doing his own self-sorts; the level at which the diagnostic material provides information about O is much further--"deeper", if you will--from O's sorting level. Consequently, there might be some bias in favor of more accurate prediction based on the interview material inherent in the experimental design. By using items which may tend to favor prediction from the diagnostic material, the intent was to cancel out some of this "interview-bias". Whether or not this decision achieved the desired result cannot be adequately determined; therefore, the original source of the trait population remains as a bias in this investigation.

The limitations discussed in this section require that interpretations of this investigation be made with caution.

GENERAL IMPLICATIONS AND SUGGESTED FURTHER RESEARCH

The major finding of this study was the decrease in predictive accuracy which occurred when the Judges were required to make a second prediction of O's self-sort (this was particularly evident when the two predictions were based on different types of stimulus material). This result is a confirmation of findings in other studies (20,54) and it is quite provocative.[1] It seems to contradict the truism: the more you know about a person the better you understand him.

The data of this study suggests the "more" in the above adage requires further elaboration. One might raise the point whether the "more information" implied above should be read "more of the same kind of information". That is, hearing ten therapy interviews might lead to more accurate prediction than listening to one or two sessions; similarly, predicting from a battery of ten projective tests might lead to more accuracy than predictions based on one or two tests.

These consideration could be subjected to an empirical test by requiring Judges to make a succession of predictions (at least more than the number required in the present study)

---

[1] One interesting area of applicability would be the results obtained by Kelley & Fiske (60).

while the amount and kind of information provided was controlled. If the above speculations are valid, the successive predictions when the type of material is kept the same (all interview or all diagnostic) should become increasingly more accurate; whereas, successive predictions based on, say, the following sequence $I_1, D_1, I_2, D_2, I_3, D_3$, etc., should result in increasingly poorer predictive accuracy.

The fact that for all Judges the second prediction was made with the same statements (as used in self-sorts) suggests several factors which may be contributing to the obtained decrement in the second prediction session. Reusing the same items under different instructions could be a determining factor. This explanation, while pertinent to the present study, would not be supported by the results obtained in those investigations in which the same population of statements was used for all the sortings. Unless one is willing to posit a cumulative effect in which each successive sort becomes increasingly inaccurate due to the effects of the preceeding sorts, such an explanation does not seem tenable.

It would be possible to test the appropriateness of the above position with the following experimental design: Two groups of Judges attempt to predict the self-sort of an O from the same stimulus material. Group 1 uses the same Q-statements for all the sortings (self- and predictive-sorts); Group 2 uses a different trait-population for each sorting.

If there is a "building-up of interference" then Group 1 should show a decrement on the second prediction and Group 2 should not show this effect.

Another possible consequence of using the same items for both the self-sort and the second predictive-sort involves the issue of threat. It may be that a Judge who "sees" the O as being severely disturbed, attempts to sort the items on the second predictive task differently than he sorted these same items for himself. (The fact that 7 of the 16 Js who had an RS significantly different from zero at the .05 level or better, did not have a decrease in their ACC scores on the second prediction, tends to vitiate this arguement.)

There is still another possibility, however, stemming from the fact that there tends to be a fairly high correlation between the self-sorts of the Judges and that of the O (median RS is +.439). If a Judge was trying to predict O as being different from himself, then, conceivably, his predictive accuracy might diminish.

Two methodological "wrinkles" if added to the present experimental design, would permit an empirical assessment of the above speculation. Keeping all other procedures the same, half of the Js in each experimental condition could have been told that O was a patient in an out-patient clinic; the other half of the Judges in each condition could be told that O was a student. If the factor which is responsible for the de-

crease in accuracy is the threat involved in predicting a disturbed person as similar to oneself, then the effect should be greater in the first group.

A more subtle procedure would be to have half of the Js do their self-sorts first (as in this study) and half of them do the self-sort after the second prediction. If RS for the first group is consistantly higher than for the second one, the importance of threat as a variable would have been demonstrated. That is, the second group of Judges would have systematically attempted to "see" themselves as different than the O.

It would be possible to combine the two above mentioned procedures into one experimental design. This would permit the measurement of the effectiveness of the different instructions. Presumably, this could result in a measure of credulity; a factor possibly influencing the obtained results.

A final factor which may be contributing to the decrement in ACC from session 'one' to session 'two', and which has important (though rarely emphasized) implications for diagnostic and psychotherapeutic theory, is the role played by feed-back.[2]

In the course of this investigation it became clear, particularly in the I"D and D I" conditions, that the Judges

---

[2] The implications of the role played by feed-back in regard to diagnosis and therapy were pointed out by Dr. D.L.Grummon.

were "weighing" the information acquired from the two stimulus sources with which they were provided. Several Judges, for example, commented under the I"D condition: "The tests make him look 'sicker' than I thought he was from the interview material"; when I" followed D, a Judge might say: "I guess he's in better shape than I thought he was from the tests".

These comments indicate that the second stimulus material was being evaluated in terms of the Judge's initial impressions of O (gained from the first stimulus material). Yet the Judges were unaware of the accuracy of their impressions; of the 16 Js whose first ACC score was significantly different from zero at the .05 level or better, only two reached this level of significance on the second prediction. The Judges' inability to assess how accurate their impressions were could be a factor in the decrease in predictive accuracy.

On the basis of the Judges' comments, it would appear that the feed-back they received--the information contained in the second stimulus material--was misinterpreted. Almost invariably a Judge would "correct" his initial ideas in terms of the cues he perceived in the second stimulus. In this effort to gain greater accuracy, the Judge revised his initial impressions (which unknown to him were relatively accurate) and, as a consequence, he became less accurate.

It would appear, then, that the Judges in this study were operating with the erroneous conviction that "more in-

formation" was equivalent to "more accurate information".

In highlighting the role played by feed-back in this study, it is felt that perhaps a key concept in the differentiation between a diagnostic and a therapeutic approach has been brought into focus. In therapy sessions one has continuous feed-back; in diagnostic testing one does not.[3] The feed-back in therapy enables the therapist to continuously correct his impressions of the client. In this way the therapist attains an ever increasing, more accurate understanding of his client's thoughts, feelings, loves, hates, etc.

Diagnosis, on the other hand, is a system of one-way communication. The diagnostician has almost no feed-back on the impressions he obtains from the diagnostic material.

Thus it would seem, that feed-back in the therapy interviews provides the therapist with a type of learning situation. That is, the therapist is able to test his hypotheses regarding the client continuously, retaining those which receive confirmation-- depending on the theoretical orientation of the therapist--and discarding those which do not. One wonders whether approaching the therapeutic interaction in these terms does not provide an avenue for the application of many of the concepts of learning theory to this area of clinical psychology.

---

3 Bruner's formulation assumes a continuous feed-back process. Since in this study this condition did not obtain, the results of this investigation cannot be considered a crucial test of his position. Rather, as stated previously, Bruner's theories served as the theoretical framework for this research.

The concept of feed-back would appear to have still a further application to Bruner's formulations. It is suggested that the process which Bruner terms "a closer look" is, in essence, feed-back; while gating results from no feed-back or the "wrong kind" of feed-back. In other words, where the feed-back and the individual's categorization system are congruent--of the same kind or at the same level of awareness--the result is more veridical perception as in the "closer look". However, when there is incongruence between feed-back and categorization system, gating results and the ability to profit from additional stimulus cues is diminished.

To translate these speculations into the terms of the present study, the experimental design used provided the Js with only a limited amount of feed-back. The Judges were able to assess their impressions in terms of additional information, but they did not have any objective criterion on which to based this assessment. In the heterogeneous conditions (I"D and D I") the second stimulus material--the additional material or feed-back--was different than that used by the Judges to gain their initial impressions. Thus, in the heterogeneous experimental conditions an incongruence existed between the initial categorization system and the feed-back available to the Judges. As a consequence, the feed-back appears to have been misinterpreted; the Judges altered their initial impressions resulting in less accuracy.

In view of these speculations (and the results obtained in this study) it would appear that the usefulness of pre-therapy formulations is open to serious questioning. The data appear to lend support to Rogers' (87) reluctant conclusion that diagnosis as a necessary prelude to psychotherapy may be a "protective alternative to the admission that it is, for the most part, a colossal waste of time." It may well be that the vast amount of time spent, prior to beginning a psychotherapeutic relationship, in formulating plans, goals, dynamic characterological descriptions, etc., could be more fruitfully utilized.

The results of this investigation have implications not only directly related to the specific issues to which the study was addressed, but ramifications in other areas as well.

This research employed a methodology which might easily lend itself as a method for pairing therapists and clients. Given a recorded intake interview and the self-sort of a prospective client, the selection of a therapist could be made on the basis of which of the available therapists could most accurately predict this individual's self-sort. Indeed, these ACC scores could be the independent variable and the success of the therapy the dependent variable in such an investigation. The ACC scores could be varied, unknown to the therapists, through a careful pairing of therapist and O.

Although it was not the major focus of this study, the failure to replicate the clearcut relationship between level of experience and accuracy of prediction reported by Kell (59), suggests that a more intense investigation of experience as a mediating variable in this type of research should be done.

The methodological innovation introduced in this study of using two equivalent populations of Q-statements (containing clusters of items) appears to be justified; particularly in regard to the increase in experimental control it engenders. The data of this study could be utilized in the development of a well standardized, reliable, and internally consistent (insofar as the clusters are concerned) trait population for use in future research involving Q-methodology.

Through an analysis of the self-sorts of the clinicians, hypotheses for additional research could be generated. For example: what are the characteristics of the "good predictor"?; What is the realtionship between "diagnostic" or "empathic" orientation and experience?; Which orientation is associated with a higher degree of successful outcomes in therapy?

The general focus of these suggested research projects would be an attempt to isolate factors, both in the therapist and in the therapeutic relationship, which would be highly related with successful psychotherapeutic interactions.

SUMMARY AND CONCLUSIONS

This investigation attempted to provide empirical data pertinent to both a practical problem and a theoretical formulation. The former concerned the usefulness of formal diagnostic procedures in therapy; the latter dealt with a hypothetical construct--gating--introduced by Bruner in a theory relating to perceptual phenomena.

To the extent that the ability to predict a client's self-sort--being aware of a client's phenomenological world --is deemed an important aspect of psychotherapy; and, to the extent that Q-methodology provides an adequate, quantifiable method for ascertaining a client's self-concept, the results of this study appear to have pertinence for the above mentioned practical problem.

The data in this study fail to confirm either the "diagnostic" or the "a-diagnostic" viewpoints. It was found that utilization of formal diagnostic material neither aids nor hampers the ability to accurately predict a client's self-sort. That is, clinicians predict equally well whether or not diagnostic material is made available to them. However, if both diagnostic material and interview material are provided, and predictions are made on the basis of both types of material (regardless of the sequence of presentation) clinicians tend to predict worse on the second occasion.

These results were interpreted as indicating that the categorization system the Judges develop (based on the first stimulus material) "interferes" with the full utilization of the information provided by the second stimulus material.

The above statement is, in essence, a rewording of the gating principle which provided the model upon which this study was based. The general proposition states:

> the type of preconceived categorization system with which one approaches events will affect the sensitivity to nuances in those events

Support for the validity of the proposition would obtain from the verification of two more specific hypotheses. Hypothesis 1 stated:

> the sensitivity of an individual to nuances of a situation will be diminished if he approaches that situation with a preconceived categorization system

In the main, the data of this study supported this contention. It was found that, regardless of the type of material employed or the sequence of presentation utilized, a marked tendency existed for clinicians to predict significantly poorer on the second of two prediction sessions.

Hypothesis 2 stated:

> the sensitivity of one individual for another person will increase as the degree of congruence between the categorization systems of the two individuals increases

The results obtained were in the general direction of supporting this hypothesis; however, they were not at the level of statistical significance.

Consequently, the findings of this investigation offer only partial support for the general proposition. Significant evidence for the existence of a type of gating mechanism was found. In an attempt to discover a factor which might mediate the functioning of a gating process--the existence of a high degree of similarity between the persons involved in the interaction--the evidence is merely suggestive.

B I B L I O G R A P H Y

1 Alexander,F & French,T.M., et al Psychoanalytic Therapy. Ronald; N.Y. 1946

2 Allport,F.H. Theories of Perception and the Concept Structure. Wiley; N.Y. 1955

3 Ash,P. "The Reliability of Psychiatric Diagnosis" J.Abn. & Soc. 1949 34:272-277

4 Baker,B.O. & Sarbin,T.R. "Differential Mediation of Social Perception as a Correlate of Social Adjustment" Sociometry 1956 xix:69-83

5 Bellak,L., Pasquerelli,B.A., & Braverman,S. "The Use of the TAT in Psychotherapy" J.Nerv. & Ment. Dis. 1949 110:51-65

6 Berg,I. "Response Bias and Personality: The Deviation Hypothesis" J.Psychol. 1955 40:61-72

7 Bender,I.E., & Hastorf,A.H. "On Measuring Gereralized Empathic Ability (Social Sensitivity)" J.Abn. & Soc. 1953 48:503-6

8 Blake,R.R & Vanderplas,J.M. "The Effect of Pre-recognition Hypotheses on Veridical Recognition Thresholds in Auditory Perception" J.Person. 1950 xix:95-115

9 Block,J.A. "A Comparison of the Forced and Unforced Q-sorting Prodedures" J.Ed.Psych.Meas. 1956 16: 481-493

10 Bordin,E.S. "Diagnosis in Counseling and Psychotherapy" J.Ed.Psych.Meas. 1946 6:169-184

11 Bruner,J.S. "On Perceptual Readiness" Psych.Rev. 1957 64:123-152

12 Bruner,J.S., Goodnow,J.J., & Austin,G.A. A Study of Thinking. Wiley; N.Y. 1956

13 Bruner,J.S. & Postman,L. "On the Perception of Incongruity" J.Pers. 1949 xviii:206

14 Brunswik,E. Systematic and Representative Design of Psychological Experiments. U.Calif.Press;Berkely 1949

15 Cameron,D.E. "A Theory of Diagnosis" in Current Problems in Psychiatric Diagnosis. (Eds. Hoch,P.H. & Zubin,J.) Grune & Stratton; N.Y. 1953

16 Cohen,J. "An Aid in the Computation of Correlations Based on Q-sorts" Psych. Bull. 1957 54:138-9

17 Coleman, J.L. "Teaching Basic Psychotherapy" Amer. J. Ortho. 1947 17:622-7

18 Cronbach,L.J. "Correlation Between Persons as a Research Tool" in Psychotherapy: Theory and Research. (Ed. Mowrer,O.H.) Ronald; N.Y. 1953

19 Cronbach,L.J. & Gleser,G.G. "Assessing Similarities Between Profiles" Psych.Bull. 1953 50:456-472

20 Dailey,C.A. "The Effects of Premature Conclusion upon the Acquisition of Understanding a Person" J.Psych. 1952 33:133-152

21 Deabler,R.F. "The Therapeutic Use of the TAT" J.Clin. Psych. 1947 iii:246-252

22 Duncan,D.B. "Multiple Range and Multiple F Tests" Biometrics 1955 11:1-42

23 Dunnett,C.W. "A Multiple Comparison Procedure for Comparing Several Treatments with a Control" J.Amer. Stat.Assoc. 1955 50:1096-1121

24 Dymond,R.F. "A Scale for the Measurement of Empathic Ability" J.Consult.Psych. 1949 13:127-133

25 Dymond,R.F. "Personality and Empathy" J.Consult.Psych. 1950 14:343-350

26 Dymond,R.F. "Can Clinicians Predict Behavior?" J.Pers. 1953 xxii:151-161

27 Ebel,R.L. "Estimation of Reliability of Ratings" Psychometrika 1951 16:407-424

28 Edwards,A.L. Experimental Design in Psychological Research. Rinehart; N.Y. 1950

29 Edwards,A.L. "The Relationship between the Judged Desirability of a Trait and the Probability the Trait will be Endorsed" J.Appl.Psychol. 1953 37:90-3

30 Edwards,A.L. Personal Preference Schedule. Copyright:1953 Psych.Corp., N.Y.

31 Edwards,A.L. & Horst,P. "Social Desirability as a Variable in Q Technique Studies" J.Ed.Psych.Meas. 1953 13:620-5

32 Fenichel,O. Problems of Psychoanalytic Technique. Psychoanalytic Quarterly; N.Y. 1941

33 Fiedler,F.E. "A Method for Objective Quantification of Certain Counter-transference Attitudes" J. Clin.Psych. 1951 vii:101-7

34 Fiedler,F.E. "An Exploratory Study of Unconscious Feeling Reactions in 15 Patient-Therapist Pairs" J.Abn. & Soc. 1952 47:446-453

35 Fiedler,F.E. "The Psychological-Distance Dimension in Interpersonal Relationships" J.Pers. 1953 xxii:142-150

36 Fiedler,F.E. "Quantitative Studies on the Role of Therapist's Feelings towards their Patients" in Psychotherapy:Theory and Research. (Ed. Mowrer,O.H.) Ronald; N.Y. 1953

37 Fordyce,W.E. "Application of a Scale of Dependency to Concepts of Self, Ideal-Self, Mother, and Father" Unpubl. PhD Thesis, U. of Washington; 1953

38 Forer,B.R. & Tolman,R.S. "Some Characteristics of Clinical Judgment" J.Consult.Psych. 1952 16:347-352

39 Frank,G.H. "Note on the Reliability of Q-sort Data" Psychol.Repts. 1956 2:182

40 Frank,L.K. "Projective Methods for the Study of Personality" J.Psychol. 1939 8:389-412

41 Fromm-Reichmann,F. Principles of Intensive Psychotherapy. U.Chicago Press; Chicago. 1950

42 Gage,N.L. "Judging Interests from Expressive Behavior" Psych.Monogr. 1952 66:#18

43 Gage,N.L. "Accuracy of Social Perception and Effectiveness in Interpersonal Relationships" J.Pers. 1953 xxii:128-141

44 Gage,N.L. & Exline,R.V. "Social Perception and Effectiveness in Discussion Groups" Hum.Relat. 1953 6:381-396

45 Gage,N.L. & Chronbach,L.J. "Conceptual and Methodological Problems in Interpersonal Perception" Psych.Rev. 1955 62:411-420

46 Gage,N.L., Leavitt,G.S., & Stone,G.C. "The Intermediary Key in the Analysis of Interpersonal Perception" Psych.Bull. 1956 53:258-266

47 Gibby,R.G. "The Stability of Certain Rorschach Variables Under Conditions of Experimentally Induced Sets: I The Intellectual Variables" J. Proj.Tech. 1951 15:3-26

48 Gilbert,W.H. "Counseling: Therapy and Diagnosis" pp.351-378 in Annual Review of Psychology: 1952 ; Annual Reviews, Inc.;Calif.

49 Gill,M., Newman,R., & Redlich,F.C. The Initial Interview in Psychiatric Practice. International Universities Press; N.Y. 1954

50 Gollins,E.S. "Forming Impressions of Personality" J. Pers. xxiii:65-7

51 Goodnow,JJ. & Pettigrew,T.E. "Some Difficulties in Learning a Simple Pattern of Events" Paper presented at Annual Meeting of Eastern Psychological Assoc. at Atlantic City; March 1956. Quoted in (11)

52 Halpern,H.M. "Empathy, Similarity, and Self-Satisfaction" J.Consult.Psych. 1955 19:449-452

53 Hastorf,A.H. & Bender,I.E. "A Caution Respecting the Measurement of Empathic Ability" J.Abn. & Soc. 1952 47:574-6

54 Hathaway,S.R. "Clinical Intuition and Inferential Accuracy" J.Pers. 1956 xxiv;223-250

55 Henle,H. "An Experimental Investigation of Past Experience as a Determinant of Visual Form Perception" J.Exp.Psychol. 1942 30:1-22

56 Herrock,J.E. & Nagy,G. "The Relationship Between Ability to Make a Diagnosis and to Select Appropriate Remedial Procedures" J.Gen.Psychol. 1948

57 Johnson,D.M. The Psychology of Thought and Judgment. Harpers; N.Y. 1955

58 Jones E.E. & de Charms,R. "Changes in Social Perception as a Function of the Personal Relevance of Behavior" Sociometry 1957 20:75-85

59 Kell,B.L. "An Experimental Study of the Ability to Predict the Self Concept of an Individual from his Therapeutic Behavior" Unpubl. PhD Thesis; U.Of Chicago: 1951

60 Kelley.E.L. & Fiske,D.W. The Prediction of Performance in Clinical Psychology. U.of Michigan Press; Ann Arbor 1951

61 Kelly,G.A. Psychology of Personal Constructs:Vol. I. Norton; N.Y. 1955

62 King,F.W. "The Use of Drawing of the Human Figure as an Adjunct to Psychotherapy" J.Clin. Psychol. 1954 x:65-9

63 Kleet,C.J. "The Stability of the Social Desirability Scale Values in the Edwards Personal Preference Schedule" J.Consult.Psychol. 1957 21:183-5

64 Livsen,H.H. & Nicholes,T.F. "Discrimination and Reliability in Q-sort Personality Description" J.Abn.& Soc. 1956 51:159-165

65 Luchins,A.S. Mechanization in Problem Solving. Psych. Monographs 1942 54:#248

66 Luft,J. "Implicit Hypotheses and Clinical Prediction" J.Abn. & Soc. 1950 45:756-9

67 Luft.J. "Differences in Prediction Based on Hearing vs. Reading Verbatum Clinical Interviews" J. Consult.Psychol. 1951 15:115-9

68 Lyerly,S.B. "The Average Spearman Rank Correlation Coefficient" Psychometrika 1952 17:421-8

69 McClelland,D.C., Atkinson,J.W., & Clark,R.A. "The Projective Expression of Needs:III. The Effect of Ego-Involvment, Success, and Failure on Perception" J. Psychol. 1949 27:311-330

70 Mehlman,B. "The Reliability of Psychiatric Diagnosis" J.Abn. & Soc. 1952 47:577-8

71 Menninger,K.A. A Manual for Psychiatric Case Study. Grune & Stratton; N.Y. 1952

72 Meyer,M.M. "The Direct Use of Projective Techniques in Psychotherapy" J.Proj.Tech. 1951 15:263-7

73 Muench,G.A. "The Application of Diagnostic Methods to Counseling and Psychotherapy" pp.53-72 in Recent Advances in Diagnostic Testing. (Ed. Harrower, M.R.) Thomas; Springfield, Ill. 1950

74 Murphy,G. "Affect and Perceptual Learning" Psych. Rev. 1956 63:1-15

75 Murray,H.A. Explorations in Personality. Oxford University Press; N.Y. 1938

76 Notcutt,B. & Silva,A.L.M. "Knowledge of Other People" J.Abn. & Soc. 1951 46:30-8

77 Patterson,C.H. "Is Psychotherapy Dependent upon Diagnosis?" Amer.Psychol. 1948 3:155-7

78 Postman,L. & Bruner,J.S. "Perception Under Stress" Psych.Rev. 1948 55:314-323

79 Postman,L., Bruner, J.S., & Walk,R.R. "The Perception of Error" Brit.J.Psychol. 1951 p.1 ff.

80 Powdermaker,F., "Techniques for the Initial Interview; Methods of Teaching Therapy" Amer.J.Psychia. 1948 104:642-6

81 Quinn,R.D. "Psychotherapists' Expressions as an Index to the Quality of Early Therapeutic Relationships Established by Representatives of the Non-directive, Adlerian, and Psychoanalytic Schools" Unpubl.PhD Thesis; U.of Chicago 1950. Quoted in (37)

82 Rapaport,D. "Projective Techniques and the Theory of Thinking" pp.196-203 in Psychoanalysis, Psychiatry, and Psychology. (ed. R.R.Knight) International University Press; N.Y. 1954

83 Remmers, H.H. "A Quantitative Index of Socio-Psychological Empathy" Amer.J.Ortho. 1950 20:161-5

84 Rogers,C.R. Client-Centered Therapy. Houghton-Mifflin; Boston. 1951

85 Rogers,C.R. & Dymond,R.F. Psychotherapy and Personality Change. U. of Chicago Press; Chicago. 1954

86 Rogers,C.R. "The Case of Mrs. Oak" in (85).

87 Rogers, C.R. "The Necessary and Sufficient Conditions of Therapeutic Personality Change" J.Consult.Psychol. 1957 21:95-103

88 Rogers,C.R. "Significant Aspects of Client-Centered Therapy" Amer.Psychol. 1946 1:415-422

89 Saul,L.J. "The Psychoanalytic Diagnostic Interview" Psa.Q. 1957 xxvi:76-90

90 Schafer,R. Psychoanalytic Interpretation of Rorschach Testing. Grune & Stratton; N.Y. 1954

91 Siegal,S. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill; N.Y. 1956

92 Siipola,E.M. A Group Study of Some Effects of Preparatory Set. Psych.Monogr. 1935 46:#210

93 Soskin,W.S. "Bias in Postdiction from Projective Tests" J.Abn. & Soc. 1954 49:69-74

94 Stanley,J.C. "Statistical Analysis of Scores from Counter-balanced Tests" J.Exp.Ed. 1955 xxiii:187-207

95 Stephenson,W. The Study of Behavior. U.of Chicago Press; Chicago. 1953

96 Stone,G.C., Leavitt, G.S., and Gage,N.L. "Generalization of Accuracy in Predicting Standard Persons" Unpubl. Manuscript; Bureau of Educational Research, U.of Ill. 1954; quoted in (45).

97 Taft,R. "The Ability to Judge People" Psych.Bull. 1955 21:1-23

98 Tagiuri,R., Blake,R.R., and Bruner,J.S. "Some Determinents of the Perception of Positive and Negative Feelings in Others" J.Abn. & Soc. 1953 48:585-592

99 Thorne,F.C. "Directive Psychotherapy: IV. The Therapeutic Implications of the Case History" J.Clin.Psychol. 1945 1:310-330

100 Thorne,F.C. "The Clinical Method in Science" Amer.Psychol. 1947 2:159-166

101 Thorndike,R.L. Personnel Selection. Wiley; N.Y. 1949

102 Thornton,G.R. "The Significance of Rank Difference Coefficients of Correlation" Psychometrika 1943 8:211-222

103 Tryon,R.C. "Reliability and Behavior Domain Validity: Reformulation and Historical Critique" Psych.Bull. 1957 54:227-249

104 Tupes,E.C. "An Evaluation of Personality-Trait Ratings Obtained by Unstructured Assessment Interviews" Psych.Monogr. 1950 64:#317

105 Vernon,M.D. "The Function of Schemata in Perceiving" Psych.Rev. 1955 62:180-192

106 Watson,R.I. The Clinical Method In Psychology. Harpers; N.Y. 1951

107 Wert,J.E., Neidt, C.O., & Ahmann,J.S. Statistical Methods in Educational and Psychological Research. Appleton-Century-Crofts, Inc.; N.Y. 1954

108 Wolf,R., & Murray, H.A. "An Experiment in Judging Personality" J.Psychol. 1936 3:358

109 Wyatt,D.F. & Campbell,D.T. "On the Liability of Stereotypes of Hypotheses" J.Abn. & Soc. 1951 46:496-500

*****
***
*

APPENDIX A

LIST OF JUDGES

Behrle, Fredrick J.

Bell, Fredrick B.

Brice, Norman

Bonier, Richard S.

Brownfain, John J.

Cain, Albert

Capell, Martin D.

Chodorkoff, Bernard

Erlandson, Forrest L.

Freed, Griffith O.

Feuerfile, David F.

Gentry, George M.

Grater, Harry

Hyman, Marvin

Lesser, William

Maes, John

Malos, Herbert B.

Margolis, Marvin

McDonough, Joseph M.

Mikol, Bernard

Monteiro, Antonio

Papania, Ned

Pilisuk, Marc

Reisman, John M.

Roberts, Allyn F.

Silver, Albert

Silverman, Herbert

Smith, Walter O.

Stein, Jacob

Vidulich, Robert N.

Vogtman, Walter G.

Wohl, Julian

APPENDIX B

EQUATING PROCEDURES FOR SOCIAL DESIRABILITY

Edwards' Personal Preference Schedule is constructed in the following manner.

> ...the pairs of statements comprising the items of the PPS have been matched with respect to their social desirability scale values, in as far as this was possible. The statements were scaled for social desirability using the psychological scaling methods of successive intervals.... When statements from two differing variables were paired to make up an item in the PPS, the statements were selected so that the social desirability scale values would be comparable. The extent to which this matching was successful is indicated by the intraclass correlation of .85 between the social desirability scale values of the paired statements making up the items. It was anticipated that matching statements for different personality variables upon the basis of the social desirability scale values of the statements would tend to minimize social desirability as a factor influencing responses to the items. (30)

The statements used in this test are based on Murray's 15 need categories. For each need eight or nine statements of presumably differing social desirability value were employed. Every need category--i.e., the statements representative of these need categories--is paired with every other need category twice; once coming first and once presented second in the pair. The subject is required to subscribe to one statement in each pair as being true about him.

Through careful selection it was possible to assign

statements representative of the different need categories to the two Q-arrays and still preserve the balance in social desirability that had been achieved in the construction of the test.

For example, Item 6 of the PPS contrasts a statement indicative of $\bar{n}$ Achievement with a statement representative of $\bar{n}$ Deference; Item 11 contrasts the same two need categories but in the reverse order and with different statements. $\bar{n}$ Achievement (Item 6) and $\bar{n}$ Deference (Item 11) went into Q"; $\bar{n}$ Achievement (Item 11) and $\bar{n}$ Deference (Item 6) went into Q'.

This "trading off" of statements continued until each need category was represented by four items in each Q-array.

APPENDIX C

SUGGESTED INSTRUCTIONS FOR Q-SORT

Step 1: Remove the cards numbered 1 to 9 and put them to one side.

Step 2: Read the first statement. If it is descriptive of you, put it on the table at the right. If it is not descriptive of you, put it to the left on the table. If you are doubtful or not sure that the statement is descriptive of you put it directly in front of you.

Step 3: Read each statement in turn, putting them into one of the three piles. To the right you will have all the statements that you think are descriptive of you. Directly in front of you will be those statements about which you are doubtful. To your left will be all those statements that you think are not descriptive of you.

Step 4: Now arrange the cards numbered 1 to 9 in front of you with the number 1 card to the extreme left and the number 9 card to the extreme right. Think of these cards as representing a scale with the "1" card meaning the LEAST and the "9" card meaning the MOST descriptive of you.

Step 5: Take the pile of statements at your extreme left and select the one statement that is LEAST DESCRIPTIVE of you. Place that statement on top of card number 1.

Step 6: Take the pile of statements at your extreme right and select the one statement that is MOST DESCRIPTIVE of you. Place that statement on top of card number 9.

Step 7: Repeat Step 5, for card number 2, using the statements in the left pile. Do the same for card number 8, using the statements in the right pile. Note that you are to put three statements on each of these cards.

Step 8: Continue in the way described above, working from the extremes toward the middle, until you have exactly the required number of statements in each pile as that printed on the numbered cards. You will find it easier to sort the three middle piles (4,5,6) if you work in from the extremes to the middle.

Note: If you use up all the statements in either of the initial three piles before you have completed all the sorting, then select statements from the middle, or doubtful pile, to complete the exact number required by each card.

You may find it useful and convenient to put the one above the other for each of the numbered cards. In this way you can shift statements back and forth from one card to another. You may make any changes you desire. Just be sure that you have the proper number of statements on each card.

Step 9: When you have finished, put the statements for each numbered card under that card. Pick up the nine pile of statements, putting the number 1 pile on top and the number 9 pile on the bottom. Then replace the rubber band on the pack.

APPENDIX D

STATEMENTS OF Q' AND Q" ARRAYS

	Q'	Q"
ABASEMENT	I should confess the things that I have done wrong.	I feel guilty whenever I have done something I know is wrong.
	I am timid in the presence of people I regard as my superiors.	I am depressed by my inability to handle various situations.
	When things go wrong for me, I am more to blame than anyone else.	The pain and misery I have suffered has done me more good than harm.
	I am inferior in most respects.	I feel better when I give in and avoid a fight.
ACHIEVEMENT	I would like to be a recognized authority in some job, profession, or field of specialization.	I can do things better than other people.
	I am usually successful in things undertaken.	I have usually done difficult jobs well.
	I want to accomplish something of great significance.	I would like to write a great novel or play.
	I enjoy doing tasks that others recognize as requiring skill and effort.	I like to solve puzzles that other people have difficulty with.
AFFILIATION	I participate in groups.	I have many friends.
	I often write letters to my friends.	I share things with my friends.
	I am loyal to my friends.	I have strong attachments to people.
	I prefer to do things with my friends rather than by myself.	I like to do things for people.

Q^{I}	Q^{II}
AGGRESSION	
I tell other people what I think of them.	I like telling people off when I disagree with them.
I like making fun of people that I regard as stupid.	I like criticizing someone publicly.
I attack points of view that are contrary to mine.	I get revenge when someone has insulted me.
I like to read newspaper accounts of murders.	I get so angry that I throw and break things.
AUTONOMY	
I avoid situations where I am expected to do things in a conventional way.	I like to avoid responsibilities and obligations.
I like to do things my own way.	I like to be independent.
I say what I think about things.	I often criticize people who are in a position of authority.
I come and go as I want to.	I feel free to do what I want to do.
CHANGE	
I like to eat in new and strange resturants.	I like to experience novelty and change in my daily routine.
I participate in new fads and fashions.	I like to experiment with new things.
I like to move about the country and live in different places.	I like to travel.
I like to do new and different things.	I like to meet new people.
DEFERENCE	
I like to praise someone I admire.	I often tell my superiors they have done a good job.
When planning something, I get suggestions from people whose opinions I respect.	I follow instructions and do what is expected of me.
In a group, I accept the leadership of someone else.	I accept the leadership of people I admire.

Q'	Q"
I like to find out what great men have thought about problems in which I am interested.	I like to read about the lives of great men.
*DEPENDENCE	
I take pains not to incur the disapproval of others.	I give myself utterly to the happiness of someone I love.
I tend to be apologetic and don't stand up for what I know are my real feelings.	I do a great many things just to avoid criticism.
I secretly wish that I were a child again.	I feel afraid of being alone, of not being wanted.
I feel things deeply and personally.	I have a strong sense of resposibility.
I feel lost and helpless.	I feel "out of sorts" if I have to be by myself for any length of time.
DOMINANCE	
I like to be one of the leaders in the organizations and groups to which I belong.	When serving on a committee, I like to be appointed chairman.
I argue for my point of view when it is attacked by others.	I am able to persuade and influence others to do what I want.
I like to be called upon to settle arguments and disputes between others.	I like to supervise and direct the actions of other people.
When with a group of people, I make the decisions about what we are going to do.	I am regarded by others as a leader.
ENDURANCE	
I like to finish any job or task that I begin.	I like to complete a single job or task at a time before taking on others.
I stick at a problem even when it may seem as if I am not getting anywhere with it.	When I have some assignment to do, I start in and keep working on it until it is completed.
I work hard at any job I undertake.	I often stay up late in order to get a job done.
I like to put in long hours of work without being distracted.	I like to avoid being interrupted while at my work.

Q'	Q''
EXHIBITIONISM	
I often say things that are regarded as witty and clever by other people.	
I sometimes do things just to see what effect it will have on others.	
	I like to tell amusing stories and jokes at parties.
I like to ask questions which I know no one will be able to answer.	
	I tell other people about adventures and strange things that have happened to me.
I like to talk about my achievements.	
	I like to use words that other people do not know the meaning of.
HETEROSEXUALITY	
I like to become sexually excited.	
	I like to be the center of attention of a group.
I engage in social activities with persons of the opposite sex.	
	I am regarded as physically attractive by those of the opposite sex.
I like to be in love with someone of the opposite sex.	
	I often go out with attractive people of the opposite sex.
	I like to kiss attractive persons of the opposite sex.
I often listen to or tell jokes in which sex plays a major part.	
	I participate in discussions about sex and sexual activities.
*INDEPENDENCE	
I have considerable confidence in my ability to do most of the things I undertake.	
	All modesty aside, I feel I am a pretty capable and competent person.
I disregard regulations that hamper my freedom.	
	Frankly, I feel that I don't care to go out of my way for other people.
I hesitate to consult anyone for help.	
	I almost always try to avoid people's pity.
I find it easy to turn down unreasonable requests.	
	I refuse to take suggestions from others out of pride.
My acquaintances have a rather low standard of achievment.	
	I feel I have enough friends; if anything, too many.
INTRACEPTION	
I try to understand how people feel about various problems they have to face.	
	I often put myself in someone's place and to imagine how I would feel in the same situation.

Q'	Q''
I think about the personalities of my friends and try to figure out what makes them as they are.	I like to predict how my friends will act in various situations.
I like to analyze my own motives and feelings.	I judge people by why they do something--not by what they actually do.
I like to analyze the feelings and motives of others.	I like to study and to analyze the behavior of others.
NURTURANCE	
I show a great deal of affection toward my friends.	I readily forgive people who may sometimes hurt me.
I treat people with kindness and sympathy.	I like to help other people.
My friends confide in me and tell me their troubles.	I sympathize with people when they are hurt or sick.
I like to do small favors for people.	I am generous to my friends.
ORDERLINESS	
I have my life so arranged that it runs smoothly and without much change in plans.	I have my meals organized and a definite time set aside for eating.
I plan and organize any work that I undertake.	I have my work organized and planned before beginning it.
I keep my things neat and orderly on my desk and work space.	I keep my letters and bills, and other papers neatly arranged and filed according to a system.
Any written work that I do is precise, neat , and well organized.	I like to make plans before starting in to do something difficult.
SUCCORANCE	
I expect my friends to do many small favors for me cheerfully.	I expect my friends to treat me kindly.

Q'

I like people to be sympathetic and understanding when I have problems.

I like people to feel sorry for me when I am sick.

I like people to help me when I am in trouble.

Q"

I like people to sympathize with me and cheer me up when I am depressed.

I like my friends to make a fuss over me when I am hurt or sick.

I like my friends to show a great deal of affection toward me.

APPENDIX E

RATING SCALE

(ABASEMENT)

(ACHIEVEMENT)

(AFFILIATION)

(AGGRESSION)

(AUTONOMY)

(CHANGE)

(DEFERENCE)

(DOMINANCE)

1 2 3 4 5 6 7 8 9
LEAST LIKE **(ENDURANCE)** MOST LIKE

To attack, overcome, to excel, to win, to fight, to work hard, to persist, to endure. To ride, swim, play, throw, hunt, shoot, box, wrestle, participate, compete. Athletic, energetic, muscular, competitive.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(EXHIBITION)** MOST LIKE

Performance of feelings, fantasies, speculations, actions. Attention, imaginative, subjective, human critical, hysterical, sensitive, individualistic, seductive, penetrating.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(INTRACEPTION)** MOST LIKE

To nourish, aid or protect a helpless person. To express sympathy. To 'mother' a child. Compassionate, gentle, protective, maternal, benevolent, humanitarian, merciful, tolerant, forgiving, tolerant.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(NUTURANCE)** MOST LIKE

To arrange, organize, put away objects. To be tidy and clean. To be scrupulously precise. Perfectionistic.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(ORDER)** MOST LIKE

To form an further an erotic relationship. Erotic, sensual, seductive.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(HETEROSEXUALITY)** MOST LIKE

To seek aid, protection, or sympathy. To cry for help. To plead for mercy. To adhere to an affectionate, nurturant parent. Helpless, forlorn, grieving, tragic, suppliant.

1 2 3 4 5 6 7 8 9
LEAST LIKE **(SUCCORANCE)** MOST LIKE

Independency (In your own terms)

1 2 3 4 5 6 7 8 9
LEAST LIKE MOST LIKE

Dependency (In your own terms)

1 2 3 4 5 6 7 8 9
LEAST LIKE MOST LIKE

* * * * * * * * * * * * * * * * *

Write in the prediction which you believe to be most accurate:

APPENDIX F

EXPERIENCE DATA SHEET

Number of diagnostic reports you have written and supervised _____.

Please indicate, by encircling the appropriate number below, how you would rate yourself on a diagnostician-therapist continuum. In other words, do you consider yourself as being more at the diagnostician end of such a continuum (Nos. 1,2,3) or more at the therapist end of such a continuum (Nos. 7,8,9).

1 2 3 4 5 6 7 8 9

Diagnostician Therapist

Please indicate your therapeutic orientation _____.

Circle the appropriate number below indicative of your feeling regarding the importance of diagnostic procedures to conducting therapy.

1 2 3 4 5 6 7 8 9

No importance Extremely important

Circle the number in front of the statement below which most accurately describes your experience with Q-sort technique.

1 I have never heard of Q-sorts.

2 I have read relatively little of the literature on Q-sorts.

3 I have read extensively on Q-sorts but have never done one.

4 I have done one or two Q-sorts.

5 I have done several Q-sorts (more than two).

6 I have done quite a number of Q-sorts and have read much of the literature on this technique.

In the experiment, proper, you will be asked to predict the Q-sort of a patient or client on the basis of a recorded therapy hour, and on the basis of diagnostic material. Circle the number in front of the statement below which describes your best estimate of the accuracy of your predictions.

1 My predictive Q-sort based on the interview material will be more accurate than my predictive Q-sort based on the diagnostic material.

2 My predictive Q-sort based on the diagnostic material will be more accurate than my predictive Q-sort based on the interview material.

APPENDIX G

METHOD OF SELECTION OF INTERVIEW MATERIAL

Selection of the interviews used as input material for this study proceeded in the following manner: Five successive therapy sessions were tape recorded. Two of these sessions preceeded the day on which the psychological tests were administered, one was on the day the O was tested (the tests were administered two hours after this interview), and two of the recored interviews followed the testing.

In selecting the interviews it was necessary that they be clear recordings; this requirement eliminated the interview conducted on the day of testing--the recording was of poor quality.

While listening to the remaining four recordings, it became apparent that the introduction of the tape recorder had inhibited O's spontaneity during the first recorded interview; Mr. "L" made mention of this in the second recorded session. For this reason the first recorded interview was eliminated.

It had also been decided by the author that the two interviews selected should be as close to the test date as possible. This requirement was an attempt to have O "similar" on these three occassions. Further, it was

felt that one of the selected interviews should have preceeded the testing and the other followed it. Accordingly, the second recorded interview--Session 11--and the fourth recorded interview--Session 13--were selected. These interviews became I' and I", respectively,

Following the selection of the interviews, both sessions were re-recorded in their entirety with the therapist's remarks erased.

APPENDIX H

RATIONALE FOR EXCLUSION OF
ATTRIBUTIVE SIMILARITY SCORE

In the early studies done on empathy--e.g., Dymond's study in 1949 (24)--in which a "prediction methodology" was used, the accuracy of the predictor was taken as a measure of his empathic ability. Hastorf & Bender (7,53) reported that this measure, which they termed "raw empathy," was significantly related to the similarity existing between J and O. They suggest as a remedial measure, subtracting AS from ACC; the resulting score--"refined empathy"--would not be confounded by "projection".[1]

Gage & Cronbach (45) in an extensive review of the methodological problems involved in the computation of "refined empathy" scores concluded: "Clearly Bender & Hastorf did not arrive at a measure of accuracy independent of AS and RS."

Hathaway (54), moreover, questions the applicability of the term "projection" as used by Bender & Hastorf:

> [They]...have argued that likeness between a clinician's self-ratings and his prediction of the client's self-ratings can be used as a measure of projection. This arguement is not rigorous, since this kind of predictor success is not necessarily derived by inter-

---

[1] Hastorf & Bender (53) report a significant positive correlation between projection and raw empathy, and a significant negative correlation between projection and refined empathy.

> personal mechanisms.
>
> Terms for defense mechanisms all relate to affective involvement, and a number representing interpersonal similarity is not unequivocal evidence for "projection." If the word projection is to have useful meaning, it signifies a motivated inference; the inference is derived from repressed material. If a participant checks a personality item about himself identically with his prediction of a target person, projection could be involved, but it is not a parsimonious explanation. Even empathy, which is less clearly identified, assumes a "feeling into." Again, mere identity does not establish the psychological state.

In view of these considerations the ACC score was used as the measure of achievement. AS and RS were seen as possible factors capable of influencing the magnitude of the ACC score. This formulation is in accord with the thinking of Baker & Sarbin (4), who maintain that the ACC score is "probabilistically multiple-determined." However, these authors raise a crucial question regarding the "correction" of ACC for the AS factor. Baker & Sarbin ask:

> What is the rationale which justifies correcting for "projection" but not for conceptual rigidity, personal distance, parataxic distortion, ethnocentrism, or any other variable that might possibly mediate achievement? ...To assert that a general tendency to attribute one's own attitudes to others, for instance, should not be allowed to affect accuracy scores because it does not reflect "real understanding" is to remove social perception from the realm of achievement and make of it something mysterious and detached from behavior.

Consequently, ACC scores were not corrected for AS. RS, on the other hand, enters into the rationale underlying Hypothesis 2 (cf. p.22); therefore, it was retained.

APPENDIX I

EFFECT OF DIFFERENT LEVELS OF EXPERIENCE ON ACC SCORES

The following tables present the results of an analysis of the ACC data (based on cluster"scores") taking into account the different levels of therapeutic and diagnostic experience of the Js. Two seperate analyses of the data were carried out: in the first analysis the Judges' levels of therapeutic experience were considered; in the second analysis their diagnostic exerience was the focal point.

A rho correlation between therapeutic experience and diagnostic experience yields a value of r = .699 which, for N = 32, is significantly different from zero beyond the .01 level (102). Despite his high correlation, when the eight Judges in each of the four experimental conditions were seperated into a high and low group (the median was used as the cut-off point in each of the four conditions for both variables) three of the four conditions had different Js in these groups when this dichotomy was made on the basis of the two variables. Only the I"D condition had the same four Js in the high and low groups regardless of the variable used for dichotomization--therapeutic or diagnostic experience. In view of this finding an analysis was carried on both variables.

In the following tables H and L will mean High and Low group for the "experience" variable under consideration.

TABLE X

ANALYSIS OF I"D CONDITION
JUDGES DICHOTOMIZED ON BASIS OF BOTH
DIAGNOSTIC AND THERAPEUTIC EXPERIENCE*

SOURCE OF VARIATION	SUM OF SQUARES	df	MEAN SQUARE	F
Between H and L	.1786	1	.1786	9.1775
Between Js in same Group (H or L)	1.1119	6		
Total Between Js	1.2905	7		
Between Sessions	.1175	1	.1175	1.2342
Interaction: Sessions X Experience	.0063	1	.0063	---
Interaction: Pooled Js X Sessions	.5717	6	.0952	
Total Within Js	..6955	8		
TOTAL	1.9860	15		

* Dichotomization on the basis of therapeutic and diagnostic experience led to identical placement of Js; i.e, same four Js appeared in the high and low groups regardles of variable utilized for dichotomization.

The F of 9.1775 obtained for the High and Low groups, for 1 and 6 degrees of freedom, is significant at the .05 level. An examination of the mean ACC scores for these two groups reveals that the low experience group had the higher mean ACC score.

The F of 1.2342 obtained for between sessions, for 1 and 6 degrees of freedom is not significant.

TABLE XI

ANALYSIS OF D I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
THERAPEUTIC EXPERIENCE

SOURCE OF VARIATION	SUM OF SQUARES		df	MEAN SQUARE	F
Between H and L	.1922		1	.1922	1.1755
Between Js in same Group	.9810		6	.1635	
Total Between Js		1.1732	7		
Between Sessions	.1562		1	.1562	2.0206
Interaction: Sessions x Experience	.0032		1	.0032	---
Interaction: Pooled Js x Sessions	.4640		6	.0773	
Total Within Js		.6234	8		
TOTAL		1.7966	15		

Neither the F of 1.1755 for the High and Low nor the F of 2.0206 for between sessions, for 1 and 6 degrees of freedom, reach the .05 level of significance.

The results for this same group of Judges dichotomized on the basis of their diagnostic experience are presented in Table XII on the next page. The F of 1.3084 obtained for between sessions, for 1 and 6 degrees of freedom, is not significant.

TABLE XII

ANALYSIS OF D I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
DIAGNOSTIC EXPERIENCE

SOURCE OF VARIATION	SUM OF SQUARES	df	MEAN SQUARE	F
Between H and L	.0037	1	.0037	---
Between Js in same Group	1.2868	6	.2144	
Total Between Js	1.2905	7		
Between Sessions	.1175	1	.1175	1.3084
Interaction: Sessions x Experience	.0387	1	.0387	---
Interaction: Pooled Js x Sessions	.5393	6	.0898	
Total Within Js	.6695	8		
TOTAL	1.9860	15		

TABLE XIII

ANALYSIS OF I'I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
THERAPEUTIC EXPERIENCE

SOURCE OF VARIATION	SUM OF SQUARES		df		MEAN SQUARE	F
Between H and L	.0104		1		.0104	---
Between Js in same Group	.8839		6		.1473	
Total Between Js		.8943		7		
Between Sessions	.1369		1		.1369	1.5211
Interaction: Sessions x Experience	.0888		1		.0888	---
Interaction: Pooled Js x Sessions	.5404		6		.0900	
Total Within Js		.7661		8		
TOTAL		1.6604		15		

The F of 1.5211 for between sessions, for 1 and 6 degrees of freedom, is not significant.

Table XIV presents the results of an analysis of the I'I" condition dichotomized on the basis of diagnostic experience. This table is given on the following page. The Fs in Table XIV of 2.2740 for between sessions and of 4.4451 for the interaction of sessions x experience, for 1 and 6 degrees of freedom, do not reach the required level of significance.

TABLE XIV

ANALYSIS OF I'I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
DIAGNOSTIC EXPERIENCE

SOURCE OF VARIATION	SUM OF SQUARES		df		MEAN SQUARE	F
Between H and L	.0922		1		.0922	---
Between Js in same Group	.8021		6		.1336	
Total Between Js		.8943		7		
Between Sessions	.1369		1		.1369	2.2740
Interaction: Sessions x Experience	.2676		1		.2676	4.4451
Interaction: Pooled Js x Sessions	.3616		6		.0602	
Total Within Js		.7661		8		
TOTAL		1.6604		15		

TABLE XV

ANALYSIS OF I"I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
THERAPEUTIC EXPERIENCE

SOURCE OF VARIATION	SUM OF SQUARES		df		MEAN SQUARE	F
Between H and L	.0038		1		.0038	---
Between Js in same Group	.2465		6		.0410	
Total Between Js		.2503		7		
Between Sessions	.0283		1		.0283	---
Interaction: Sessions x Experience	.0000		1		.0000	---
Interaction: Pooled Js x Sessions	.7036		6		.1172	
Total Within Js		.7319		8		
TOTAL		.9822		15		

None of the Fs in Table XV are significant.

Table XVI, presented on the following page, gives the results of a second analysis on the I"I" group. For the data in this table the group was dichotomized on the basis of diagnostic experience. The obtained F in TableXIV of 3.3042 for the interaction Sessions x Experience, for 1 and 6 degrees of freedom, is not significant.

TABLE XVI

ANALYSIS OF I"I" CONDITION
JUDGES DICHOTOMIZED ON BASIS OF
DIAGNOSTIC EXPERIENCE

SOURCE OF VARIANCE	SUM OF SQUARES		df	MEAN SQUARE	F
Between H and L	.0001		1	.0001	---
Between Js in same Group	.2502		6	.0417	
Total Between Js		.2503	7		
Between Sessions	.0283		1	.0283	---
Interaction: Sessions x Experience	.2498		1	.2498	3.3042
Interaction: Pooled Js x Sessions	.4538		6	.0756	
Total Within Js		.7319	8		
TOTAL		.9822	15		

The tables on the next two pages, Table XVII and Table XVIII contain the results for all the experimental conditions considered simultaneously. Table XVII gives the results when all the Js in all the conditions are dichotomized into high and low therapy-experience groups. For Table XVIII the Js were similarly dichotomized on the basis of diagnostic experience.

TABLE XVII

ANALYSIS OF ALL CONDITIONS
JUDGES DICHOTOMIZED ON BASIS OF
THERAPEUTIC EXPERIENCE

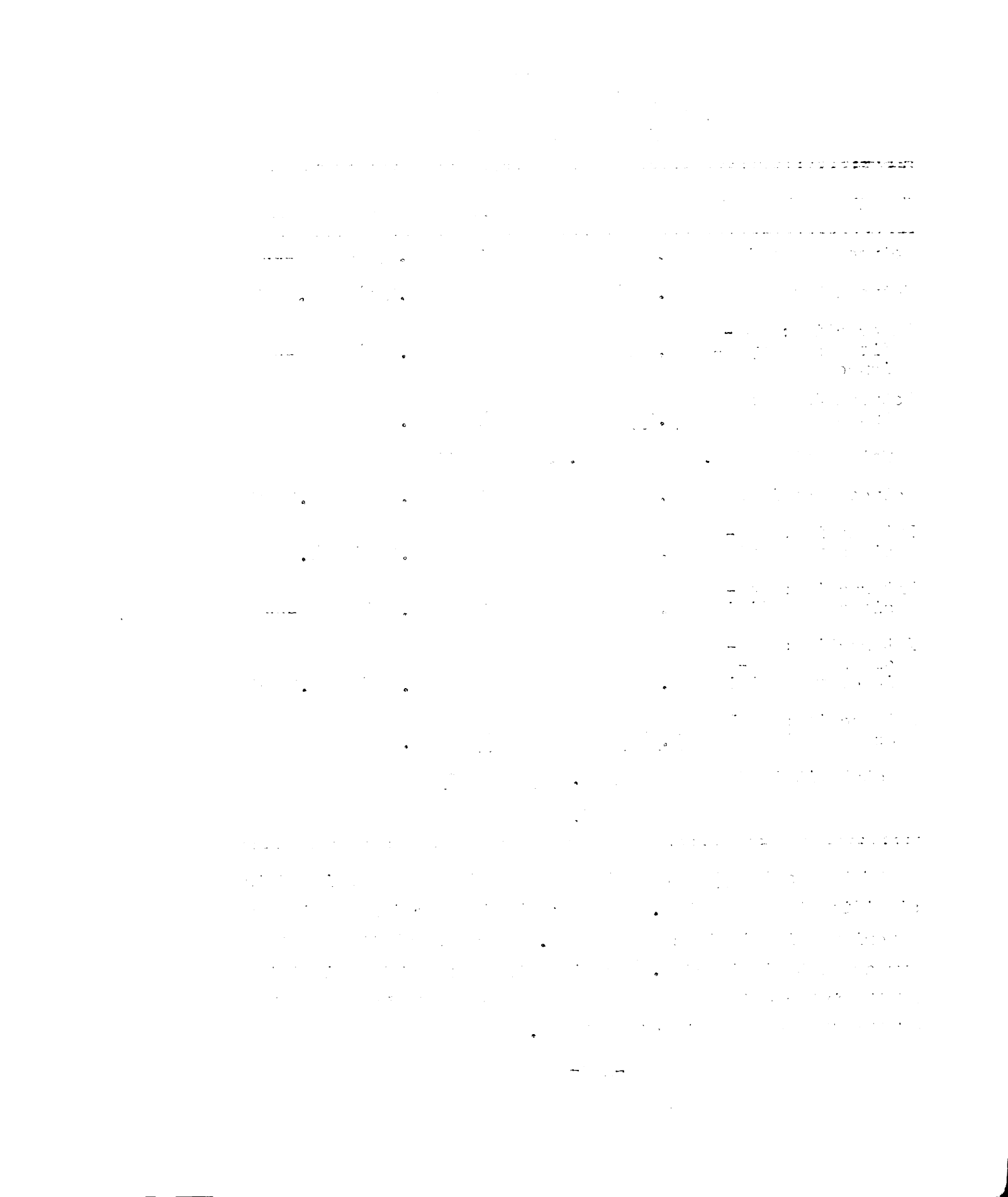
SOURCE OF VARIANCE	SUM OF SQUARES		df		MEAN SQUARE	F
Between Conditions	.2209		3		.0736	---
Between H and L	.1683		1		.1683	1.2531
Interaction: Conditions x Experience	.2167		3		.0722	---
Between Js in same Group	3.2233		24		.1343	
Total Between Js		3.8292		31		
Between Sessions	.5685		1		.5685	6.5344
Interaction: Sessions x Experience	.0188		1		.0188	---
Interaction: Sessions x Conditions	.0830		3		.0195	---
Interaction: Sessions x Experience x Conditions	.0585		3		.0195	---
Interaction: Pooled Js x Sessions	2.0880		24		.0870	
Total Within Js		2.8168		32		
TOTAL		6.6460		63		

The F of 1.2531 for H and L, for 1 and 24 degrees of freedom is not significant. The F of 6.5344 for between sessions, for 1 and 24 degrees of freedom is significant between the .05 and .01 levels.

TABLE XVIII
ANALYSIS OF ALL CONDITIONS
JUDGES DICHOTOMIZED ON BASIS OF
DIAGNOSTIC EXPERIENCE

SOURCE OF VARIANCE	SUM OF SQUARES		df		MEAN SQUARE	F
Between Conditions	.2209		3		.0736	---
Between H and L	.1524		1		.1524	1.0971
Interaction: Conditions x Experience	.1212		3		.0404	---
Between Js in same Group	3.3347		24		.1389	
Total Between Js.		3.8292		31		
Between Sessions	.5685		1		.5685	8.3357
Interaction: Sessions x Experience	.3006		1		.3006	4.4076
Interaction: Sessions x Conditions	.0809		3		.0269	---
Interaction: sessions x Experience x Conditions	.2262		3		.0754	1.1055
Interaction: Pooled Js x Sessions	1.6406		24		.0682	
Total Within Js		2.8168		32		
TOTAL		6.6460		63		

Of the tabled F values listed above "between sessions" is significant beyond the .01 level; the interaction sessions x experience is significant at the .05 level, and the remaining two are not significant. An investigation of the significance of the interaction reveals the significance to be the result of the "between sessions" variable.



APPENDIX J

ANALYSIS OF DATA: SCORES BASED ON ITEMS

The data reported in the body of this study are based on the scores attained by the Js when the clusters of items were used as the basis for the necessary computations. The data were also analyzed with the individual items used as the basis for the calculations. A rho correlation, based on the ranked mean ACC scores(for each Judge)attained by these two methods of analysis, was computed. The obtained value, r= .830 is significantly different from zero at the .01 level. Thus there would appear to be a significant degree of similarity between the way the Judges (based on mean ACC) are ranked by these two methods of analysis--clusters vs items.

Table XIX (p.105) gives the correlation scores for all the Judges, when computed on the individual items.

A comparison of this table with Table IV (p.49) reveals the differences in the obtained values of the scores resulting from the two methods of computation. When based on clusters the ACC scores range from -.217 to .733; when based on items the ACC scores range from -.102 to .449. The cluster correlations are based on 17 "need scores", the item correlations are based on 70 "item scores". Twenty of the sixty-four ACC scores reach the .05 level of significance or beyond, in the former method; 30 do in the latter one; .468 and .234 are the respective .05 levels.

TABLE XIX

ACC AND RS CORRELATIONS
FOR ALL JUDGES
(Based on Items)

GROUP	JUDGE	SESSION ONE	ACC	SESSION TWO	GP. MEAN	RS
	A	-.033		.010		-.009
	B	.375**		.076		.300*
	C	.346**		.352**		.360**
	D	.323**		.214		.105
I"D	E	-.027		-.056		.403**
	F	.237*		.294*		.395**
	G	.254*		.386**		.289*
	H	.088		-.061		.220
	Mean --	.200*	---	.154	.179	
	I	-.033		.205		.317**
	J	.116		.013		.426**
	K	.228		.208		.243*
	L	.128		.231		.435**
D I"	M	.180		.415**		.111
	N	.369**		.306**		.269*
	O	.180		.203		.386**
	P	.421**		.386**		.438**
	Mean --	.203*	---	.249**	.226*	
	R	.254*		.154		.326**
	S	.363**		.266*		.271*
	T	.392**		.228		.231
	U	.243*		.025		.306**
I'I"	V	.220		.028		.228
	W	.254*		.418**		-.130
	X	.395**		.449**		.214
	Y	.045		.214		.323**
	Mean --	.274**	---	.228*	.251**	
	AA	.346**		-.102		.312**
	BB	.346**		.217		.360**
	CC	.203		.323**		.346**
	DD	.002		.111		.300*
I"I"	EE	.274*		.225*		.363**
	FF	.254*		-.093		.237*
	GG	.312**		.312**		.185
	HH	.142		.094		.237*
	Mean --	.237*	---	.139	.188*	
TOTAL MEANS		.211*		.229*	.194*	

* Significantly different from zero at .05 level
** Significantly different from zero at .01 level

Despite the significant correlation found between mean ACC scores as computed by the cluster and item methods, when an analysis of variance is applied to the data obtained by the latter method (as shown in Table XX) no significant results are found. Presumably, this is the consequence of the lower reliability of the ACC scores when they are based on correlation of the individual items, cf. Table VII, p. 54.

The discrepancy in the results obtained by the use of these two methods of computation, suggests that the scoring method, used in studies employing Q-technique, is a critical factor in evaluating the results obtained.

TABLE XX

ANALYSIS OF VARIANCE OF ACC SCORES*
FOR FOUR GROUPS OF JUDGES UNDER DIFFERENT
EXPERIMENTAL CONDITIONS WITH TWO SESSIONS FOR EACH GROUP

SOURCE OF VARIATION	SUM OF SQUARES		df		MEAN SQUARE	F
Between Conditions	.0600		3		.0200	.5780
Between Js in same Group	.9705		28		.0346	
Total Between Js		1.0305		31		
Between Sessions	.0213		1		.0213	1.5106
Interaction: Sessions x Conditions	.0475		3		.0158	1.1205
Interaction: Pooled Js x Sessions	.3954		28		.0141	
Total Within Js		.4642		32		
TOTAL		1.4947		63		

* Correlations based on individual items

The Fs of .5780 and 1.1205 for "between conditions" and the interaction sessions x conditions, for 3 and 28 degrees of freedom, are not significant. The "between sessions" F value of 1.5106, for 1 and 28 degrees of freedom, is not significant.

Thus none of the results obtain, when the data are based on correlations between the individual items. The statistical significance is found when correlations based on clusters of items--needs--are utilized.

*****
***
*