THE EFFECT OF VALIDATING AND INVALIDATING EVIDENCE ON CONSTRUCT ORGANIZATION

MARY PHILOMENA LAWLOR

THE EFFECT OF VALIDATING AND INVALIDATING EVIDENCE ON CONSTRUCT ORGANIZATION

by

Mary Philomena Lawlor, B.A., B.Ed.

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

Department of Psychology
Memorial University of Newfoundland

August 1977

St. John's Newfoundland

ABSTRACT

This study tested Bannister's (1963) serial invalidation hypothesis that repeated invalidation of one's judgments loosens construct organization. Sixty university students judged three sets of eight career roles on 16 different constructs. The first eight constructs were termed target constructs since they were the focus of validation and invalidation during treatment. The second eight were termed non-target constructs since no feedback was provided for judgments involving these constructs. Subjects rated eight career roles (pre-treatment), then rated eight more receiving feedback on the eight target construct judgments for each individual career role (treatment), and finally rated eight more without feedback (post-treatment). Treatment consisted of four combinations of validating and invalidating levels of both qualitative and quantitative feedback types. One group was assigned to each of the four feedback conditions. Qualitative feedback consisted of evaluative comments by the experimenter on the subjects' performance. Quantitative feedback consisted of fake ratings shown to the subjects, from which they could assess the accuracy of their performances.

Data were analysed separately for subjects with strong pre-treatment relations among constructs and those with weaker pre-treatment construct relations, since these groups tend to modify construct organization differently.

For subjects with strong construct relations the results indicated that those who received totally invalidating feedback loosened construct relations during and after treatment significantly more than those who received totally validating feedback. This result held for both target and non-target constructs. The mixed feedback groups closely resembled each other with average scores on both the strength and consistency of relations which uniformly fell between the validated and invalidated groups. However, for subjects with weaker relations among constructs, there were no significant differences on any measure or any condition. Analysis of the effect of feedback type supported an additive as opposed to an interactive contribution. The evidence from this study strongly supports the serial invalidation hypothesis, but only for subjects with strong construct relations. Implications for the study of thought disorder were discussed.

ACKNOWLEDGEMENTS

I appreciate the efforts of the many people who contributed in many ways to making this research worthwhile for me.

For direct involvement, I would like to express my gratitude to Dr. Larry Cochran, my supervisor. His thorough knowledge and formative advice provided me with the necessary theoretical understanding for this study. But mostly I appreciated his patience and encouragement at all stages of the research. I am also grateful to the other members of my committee: Dr. David Hart, who first introduced me to this area of research, and Dr. Michael Stones, whose enthusiasm for Personal Construct Theory was catching. Their critical reviews and willingness to help at all times certainly made the task enjoyable as well as interesting.

Thanks are also due to two special friends, Mildred and Judy, for their willingness to listen to and critically comment on my unpolished ideas, and for some laborious editing of the ensuing drafts of those ideas.

And finally, to my parents, thank you for your continuous support and belief in me.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

In recent years, there has been a broad convergence upon the importance of expectation (anticipation, prediction, and so on) in understanding human experience and behavior. While differing in many details, certain theorists generally assume some structuring capacity which allows organisms to establish an orderly representation of the world. This representation (e.g. schema, cognitive organization) generates expectations which can be validated or invalidated. The validational outcomes of construing are then thought to modify or stabilize a person's representation system.

Cognitive approaches have always stressed expectancy as a central construct in the explanation of behavior. One major school of cognitive psychology, for instance, emphasizes hypothesis testing. (Levine, 1975, provides an excellent history of the development of hypothesis-testing theory.) Early investigators in animal learning and behavior (Hamilton, 1911; Yerkes, 1961) noted that prior to learning, animals displayed behavior which appeared to be more or less organized and systematic. In perceptual discrimination studies, Lashley (1929) suggested that these systematic responses prior to learning appeared to represent attempted solutions to problem tasks. Krechevsky (1932) coined the concept of hypothesis-testing to account for 'systematic attempted solutions'

observed by Lashley and others, and to offer an alternative to the prevalent notion of trial and error learning. In their classic study of thought, Bruner, Goodnow, and Austin (1956) applied hypothesis-testing to human concept learning, suggesting that anticipated outcomes might be the link between human motivational status and judgment behavior. Restle (1960, 1961, 1962) later attempted to formalize a strategy-selection model of learning. The rate of learning was assumed to be dependent upon the number of relevant strategies available. And strategies were defined as predictions which were capable of being tested. Recently, Levine (1963, 1966, 1970; Levine, Leitenberg & Richter, 1964; Frankel, Levine & Karpf, 1970; Gholson, Levine & Phillips, 1972) has refined a highly sophisticated model of hypothesis-testing to account for human learning. Briefly, people are thought to approach a particular problem with different sets of potential expectancies, and these expectancies or hypotheses affect behavior in the problem situation in specifiable ways. If a person selects from the correct hypothesis set, his expectancies will be validated by the outcomes. Consequently, he will continue to sample from that set. However, if prediction is based upon an incorrect hypothesis set, outcomes will invalidate expectations and predictions based upon alternative hypothesis sets becomes likely. In this model, continued predictions from an incorrect hypothesis set explains instances of non-learning.

Behaviorism has been perennially divided on the issue of expectation. One of the most influential early opponents

to the radical S-R paradigm of learning was E.C. Tolman. Tolman asserted that what the organism acquires in a situation is "an expectancy, a sign gestalt, a cognitive structure, a cognitive map relative to that environment" (1949, p. 150). Tolman's position is currently reflected and refined by a number of behaviorists. For instance, Estes (1969) conducted a series of experiments with human subjects which revealed major inadequacies in the Law of Effect. Estes suggested that reinforcement does not strengthen or weaken the response that produces it. Rather, learning actually occurs as a function of contiguity, independently of the reward values of the outcomes. Bolles' expectancy model of learning was designed to account for findings that indicate a failure of reinforcement to condition certain behaviors (e.g. Breland & Breland, 1961; Brown & Jenkins, 1968; Williams & Williams, 1969). According to Bolles (1972), what is learned are "certain events, cues (S) which predict certain other important events, consequences (S*)" (p. 402). What is learned are not new responses, but the expectancy that responses produce certain outcomes. It is this set of expectancies that are assumed to differentially affect performance. Bandura (1974) has recently concluded that "so-called conditioned reactions are largely self-activated on the basis of learned expectations rather than automatically evoked" (p. 859).

Similarly, from a field as seemingly distant as, for instance, physiological psychology, expectation has assumed a central role in psychological explanations. Sokolov's

(1963) explanation of the orienting reflex assumes that the human nervous system constructs internal neuronal models of the environment. The orienting reflex is triggered by a mismatch between the expectancy generated by an internal model and the actual environmental stimuli.

From a variety of standpoints then, theorists subscribe to variations of a single general model. People develop representations of events which enable them to anticipate events. The extent to which these anticipations are validated is thought to at least partially determine the extent to which a person will feel compelled to change his representation of events or his representational system. The present study is concerned with a test of one aspect of this general position, the extent to which validating and invalidating feedback differentially change the functioning of a person's representational system.

This study is based directly upon Personal Construct Theory (Kelly, 1955). Probably no theory of personality has so thoroughly and so consistently been constructed upon the notion of anticipation. Using the analogy of man as scientist, Kelly's Fundamental Postulate states that "a person's processes are psychologically channelled by the way he anticipates events" (1955, Vol. 1, p. 46). Man virtually lives in anticipation. Like a scientist, people are assumed to form models of events, develop hypotheses or anticipations, test hypotheses through experience and action, and to evaluate results.

The representational model in this theory is a personal construct system. A set of hierarchically organized constructs (bipolar concepts such as warm/cold) are thought to supply reference axes which allow the elements (i.e. anything capable of being construed, e.g. people, objects, etc.) of the world to be ordered, to be placed within a coherent system. The basic unit of this system is the personal construct which simultaneously integrates one set of elements while differentiating them from another set. As the notion of a system suggests, at least most constructs must be related and consistently related to function manageably and coherently. That is, if individual constructs are to function together (in coordination), there must be regularity not only in the way elements are discriminated on constructs, but regularity among the constructs or dimensions of judgment. For example, if the construct 'warm/cold' is to supply a basis for making useful discriminations, it must polarize elements regularly as warm or cold. Similarly, if the commonly associated constructs of 'warm/cold' and 'kind/cruel' are to function together, they must regularly discriminate between elements, and discriminate such that 'warm' elements are consistently judged as 'kind' and 'cold' elements are consistently judged as 'cruel'.

Within this theoretical framework, the hypothesis that validational evidence affects the strength and consistency of construct relations takes on considerable importance. For if regularity within a construct system is

affected by validational evidence, then feedback concerning one's anticipations can either enable or disenable a person to construe coherently.

Kelly (1955) argued that the state of a construct system is dependent upon a person's experience with validational evidence. Later in elaborating this notion, Bannister (1960, 1962) demonstrated the importance of two state-values of a construct system, the strength and the consistency of construct relations. When a number of elements have been ranked or rated on two constructs, the strength of the relation between those two constructs is measured by a correlation coefficient. When two constructs have been correlated on two different sets of elements (or the same set at different times), the consistency of the relation can be assessed by the similarity in magnitude and direction of the correlations. Using a repertory grid technique (Bannister & Mair, 1968), Bannister found that thought-disordered schizophrenic patients could be reliably distinguished from other groups (e.g. normals) by the weakness and instability of construct relations. To account for this lack of regularity within the personal construct systems of thought-disordered schizophrenic patients, which has since been replicated by numerous investigators (Radley, 1974, summarizes most of the evidence), Bannister posed the serial invalidation hypothesis. People who have had their judgments validated will tighten (i.e. strengthen and stabilize) their construct organizations while people who have had their judgments invalidated will loosen (i.e. weaken

and vary) their construct organizations. To avoid continual invalidation, judges are thought to loosen their system of construing in order to increase the probability of validation. By weakening and varying relations, judges define their expectations more vaguely and fragmentally, thereby increasing the probability of being right in some aspects. In contrast, judges who have been continually validated in their judgments tighten construction in order to define expectations more specifically and coherently.

In a series of four experiments, Bannister (1963, 1965) put this rationale to the test. All four experiments were variants of one experimental paradigm. Normal subjects were presented with photographs, informed that the task assessed their ability to judge individuals on the basis of facial appearance, and asked to rank order the photographs of persons on a set of provided adjectives. Each experiment was conducted over a five to ten day period with one to two trials per day. In the first and fourth experiments, subjects were validated on some constructs and invalidated on others. In the second and third experiments, subjects were either validated, invalidated, or uninformed on all construct judgments.

For each trial, the rank ordering of photographs on each construct was correlated with the rank orderings on every other construct. From these rank order correlations, two measures were computed. First, the intensity of construct relations was measured by squaring, multiplying by one hundred,

and summing the correlations for each trial. For those experiments involving both validation and invalidation for each subject, correlations were divided into within-validation, within-invalidation, and between-set scores. Second, the total number of reversals in correlations were added. A reversal is a shift from a significant relationship between two constructs to a significant relationship in the opposite direction (i.e. from positive to negative or vice versa). A high reversal score would then indicate a considerable amount of inconsistency in the way constructs are related. For example, good might correlate positively and significantly to sincere on the first trial, and correlate negatively and significantly on the second trial.

Results from the first three experiments were quite similar. Validation significantly increased the strength of correlations among validated constructs. When their judgments were validated, subjects tended to tighten the relationships among constructs. Second, except for the second experiment, invalidation produced a significant increase in reversal scores. That is, when their judgments were invalidated, subjects tended to vary the way constructs were related. However, despite the number of radical changes in the direction of construct relations, invalidation did not lead to a significant decrease in correlational strength.

In the fourth experiment, Bannister attempted to correct two possible flaws in his experimental design. First, if a subject is invalidated on all constructs, he is apt to

be faced with chaos. The total loss of structure might be too great a threat, leading subjects to preserve relations at any cost (namely at the expense of accuracy). To eliminate this possibility, subjects were validated on one set of constructs and invalidated on another set of constructs. Second, a construct can be validated indirectly through its relations. For example, if a person were validated on his use of 'good', it could also validate a closely related construct such as possibly 'sincere'. To eliminate this possibility, Bannister selected two sets of constructs which internally were highly related, but externally (between-set relations) were less highly related. With these refinements, validation significantly increased while invalidation significantly decreased the intensity of construct relations.

Since they provide the direction for the present study, two aspects of this series of experiments should be noted. First, the most reliable finding of these experiments was that validation strengthened construct relations. Yet as important as this result is, it has been curiously neglected. The evidence bearing upon the effect of invalidation has been over-emphasized at the expense of the solid finding that validation modifies construct organization in a particular way. Second, the effect of invalidation upon construct organization must be regarded as tentative and as it stands, unreliable. Considering the importance of this series of experiments and the amount of research it has generated (Radley, 1974, summarizes most of this work), it

is surprising that the effect of invalidation in the fourth experiment has not been replicated. Bannister's studies were exploratory. There was no pretense of completeness or adequacy. Yet the refinements and extensions which would solidly demonstrate and elaborate the way validation and invalidation modify construct organization are absent. Cochran (1973, 1976, in press) has extended Bannister's efforts in one direction, using contradictory attributes to disrupt construct organization (i.e. subjects made character judgments about a person upon the basis of attributes which were inconsistent such as 'warm' and 'inconsiderate'). While the use of contradictory attributes is certainly invalidating, it was different than the form of invalidation which Bannister used, simply showing people that their judgments were wrong. The purpose of the present experiment then is to replicate Bannister's studies directly, and to refine the use of different types of validational evidence.

Refinements

Four types of refinements on Bannister's basic paradigm were added to the present experiment. First, Bannister used (and perhaps confounded) two quite different types of feedback. One branch of study within information theory (e.g. Annett, 1969) indicates that feedback presented by the experimenter, informing the subject that he is right or wrong, influences task performance differently than feedback presented in a more objective manner (for instance,

through score cards, counters, and so on). Verbal feedback from the experimenter will be termed qualitative feedback. Objective feedback provided by score cards and the like will be termed quantitative feedback. Trowbridge and Cason (1932), for instance, demonstrated that precise directional feedback produced superior performance compared to feedback which simply informed the subject of whether his response was correct or incorrect. While the objective of this experiment and others (e.g. Smode, 1958) are quite different than Bannister's, they strongly suggest that type of feedback might influence the extent to which construct organization is modified. In the first two experiments of Bannister's series, both qualitative and quantitative feedback were used, while in the second two, only quantitative feedback was presented. To assess the impact of these two different types of feedback, the present study systematically combined qualitative and quantitative feedback. One group received validating qualitative and validating quantitative feedback. A second group received validating qualitative and invalidating quantitative feedback. A third group received invalidating qualitative and validating quantitative feedback. And a fourth group received invalidating qualitative and invalidating quantitative feedback.

Second, certain information theory research also indicates that the frequency of feedback presentation can affect task performance. For example, Lincoln (1954) found that subjects were able to learn a motor task more quickly

when feedback followed each trial rather than when summary scores followed a series of trials. Bilodeau and Bilodeau (1958) concluded that task performance is superior when feedback or knowledge of results follows each trial. Again, while the objective of these studies diverges from Bannister's studies, they suggest that the frequency of feedback might affect the extent to which construct organization is modified. In Bannister's first two experiments, qualitative feedback was presented after each trial while quantitative feedback was presented at intervals during trials. In the last two experiments, quantitative feedback was presented after each trial. To control for frequency effects, the present study used both qualitative and quantitative feedback following each set of ratings. For example, in this study, subjects rated career roles on a set of 16 constructs. After a subject rated one career role, he received fake scores showing him exactly how divergent his ratings were from the way people presumably are who assume this role. Then the subject received verbal feedback from the experimenter to the effect that his judgments were reasonably accurate or inaccurate.

Third, in Bannister's studies, there was a relatively large time span between experimental sessions. This time between sessions could have affected the impact of treatment. For example, if a person were invalidated on his use of certain constructs during experimental sessions, he would have much more time outside of these sessions to validate or re-validate his use of those constructs. Access to feedback

from the natural environment is likely to take precedence over feedback from a contrived setting. It might mask or dampen the effect of experimental invalidation and stabilize the effects of experimental validation. In the present study there was less of a temporal gap. Subjects rated eight career roles, one after the other, receiving feedback immediately after each role was rated.

Last, Bannister's use of photographs of persons is vulnerable to two types of difficulties. First, at least some subjects might not take feedback about their judgments of mere photographs seriously. That is, their failures and successes could be discounted as unimportant, trivial, and the like, since these would not be of importance in their lives. Second, it is difficult to generalize from photo-persons to real persons in the subjects' interpersonal worlds. To minimize these difficulties, the present study used career roles. Particularly for college students who are faced with career decisions, the accuracy of their perceptions of career roles is apt to be taken seriously. Also, career roles are a reality. People form, at least stereotypical, conceptions of career roles. The persons who occupy the positions of, for instance, lawyer, actor, physician, executive, and so on, are expected to have similar types of personalities, to have salient personality characteristics which fit them to the various roles. Career roles are more likely than photographs to be 'real' elements in subjects' worlds.

Hypotheses

A brief outline of the experimental tasks will serve to clarify the hypotheses. Subjects rated 24 different career roles on 16 personality constructs adapted from the 16 PF questionnaire. These ratings were divided into three different phases:

1) Pre-treatment test: Subjects rated eight roles, one at a time, on all 16 constructs.

2) Treatment: Subjects rated eight more roles, one at a time, on all 16 constructs. After each role rating, subjects received both qualitative and quantitative feedback about the accuracy of their judgments on eight constructs. These constructs on which feedback was provided were termed target constructs. No feedback was given about the remaining eight constructs. These constructs for which no feedback was provided were termed non-target constructs.

3) Post-treatment test: Subjects rated eight more roles, one at a time, on all 16 constructs.

The present study was designed to test Bannister's serial invalidation hypothesis. Invalidational evidence is expected to require subjects to weaken and vary the relations among constructs, while validational evidence is expected to encourage subjects to strengthen and stabilize the relations among constructs. Also, the design allowed tests of the extent to which these effects generalized to related constructs (non-target) and to new roles. The key test of the serial invalidation hypothesis is between the target construct grid

of the pre-treatment test and the target construct grid from the treatment phase. The other grids allow for tests of generalization. For clarification, the following diagram illustrates the different grids upon which the following hypotheses are based.

Pre-treatment test	Treatment test	Post-treatment test
Target Grid One	Target Grid Two	Target Grid Three
Non-target Grid One	Non-target Grid Two	Non-target Grid Three

The following hypotheses were tested:

1) During treatment, feedback significantly affects the intensity of target construct relations, with validating evidence strengthening and invalidating evidence weakening the correlations among constructs.

2) During treatment, feedback significantly affects the intensity of non-target construct relations, with validating evidence strengthening and invalidating evidence weakening the correlations among constructs.

3) Following treatment (when judging new roles on the post-treatment test), feedback significantly affects the intensity of target construct relations, with validating treatment strengthening and invalidating treatment weakening the correlations among constructs.

4) Following treatment, feedback significantly affects the intensity of non-target construct relations, with validating treatment strengthening and invalidating treatment weakening the correlations among constructs.

In summary, hypothesis one tests change between target grid one and target grid two, hypothesis two between non-target grid one and non-target grid two, hypothesis three between target grid one and target grid three, and hypothesis four between non-target grid one and non-target grid three.

5) During treatment, feedback significantly affects the consistency of relations among target constructs, with validating evidence stabilizing and invalidating evidence altering the pattern of relations among constructs.

6) During treatment, feedback significantly affects the consistency of relations among non-target constructs, with validating evidence stabilizing and invalidating evidence altering the pattern of relations among constructs.

7) Following treatment, (when judging new roles on the post-treatment test), feedback significantly affects the consistency of relations among target constructs, with validating treatment stabilizing and invalidating treatment altering the pattern of relations among constructs.

8) Following treatment, feedback will significantly affect the consistency of relations among non-target constructs, with validating treatment stabilizing and invalidating treatment altering the pattern of relations among constructs.

In summary, hypotheses five through eight test change between the same grids indicated in the summary of hypotheses one through four, although the focus is upon the consistency rather than the intensity of construct relations.

Last, there were two types of feedback in this study. Subjects were presented with fake score-sheets indicating how much their judgments diverged from actuality. After the subjects had a chance to consider the accuracy of their judgments for a moment, they received verbal commentary from the experimenter to the effect that they were relatively accurate or inaccurate. The first type of feedback (quantitative) is more objective (at least from the perspective of the subject) while the second type of feedback (qualitative) is more evaluative. Lacking a theoretical rationale for why one type of feedback would have more impact than another, it is assumed that each contributes independently to the modification or stabilization of construct organization. That is, using a two-way ANOVA with qualitative feedback constituting one factor and quantitative feedback the other factor, effects are expected to be additive rather than interactive.

9) In hypotheses one through four, qualitative and quantitative feedback are expected to contribute independently or additively to the effects of validation and invalidation.

10) In hypotheses five through eight, qualitative and quantitative feedback are expected to contribute independently or additively to the effects of validation and invalidation.

In summary, both the change in intensity scores and consistency scores result from an independent contribution from two types of feedback, qualitative and quantitative.

CHAPTER II

METHOD

Subjects

Sixty (31 male, 29 female) paid undergraduate students, ages 16-47 years ($\bar{X}$ = 19.5 years; S.D. = 4.77), at Memorial University of Newfoundland participated in this study. Subjects were randomly assigned to four equal groups prior to the study. This occurred by the experimenter placing cards, one square inch in area, numbered individually from one to 60, in a large box and drawing from them. The first 15 numbers were placed in Group 1, the second 15 numbers went into Group 2, etc. until all sixty numbers were in four equal groups. The numbers assigned were taken to represent subjects in order of their appearance for the study. All subjects were volunteers obtained from either the psychology department subject pool or response to on-campus advertisements for subjects.

Materials

Rhetorical aids. To strengthen belief and involvement in the experimental procedures, subjects were introduced to three different resources prior to the experiment: the 16 PF Handbook (Cattell, Eber & Tatsuoka, 1970), a 16 PF questionnaire with an answer sheet which has the 16 scales on the back, and a copy of Psychology Today (July, 1973)

which features an article by Cattell. Subjects were informed that the 16 PF is a standard and powerful way to assess personality, and that numerous studies have demonstrated that people occupying different career roles tend to have distinctive personality characteristics. The questionnaire indicates to subjects the test through which the 16 personality characteristics were measured. Thumbing through the handbook indicates numerous profiles of different career roles. And the article indicates something of the popularity of the theory, along with a picture of a very distinguished looking scientist (i.e. Cattell himself). These materials were made available during the introductory description of the experiment in order to convince subjects that the tasks were real and important.

Constructs. The 16 constructs used to rate career roles were adapted from Cattell's 16 PF test. To insure clarity, the 16 constructs were translated, where necessary, into more familiar terms. These constructs were used because they are widely used in vocational guidance as well as personality assessment, and the available materials provided assurance that the tasks would be realistic and credible. Also, having briefly observed the scope of research conducted with the 16 PF, the subjects would be more likely to accept feedback even when it diverged perhaps strangely from their expectations.

The 16 constructs were divided into two groups of eight. The first group consisted of: reserved/warm-hearted,

serious/happy-go-lucky, easily led/assertive, conservative/experimental, trusting/suspicious, practical/imaginative, self-assured/apprehensive, and emotionally unstable/emotionally stable. The second group consisted of: slack/conscientious, concrete reasoning/abstract reasoning, timid/bold, naive/shrewd, tough-minded/tender-minded, relaxed/tense, undisciplined/controlled, and self-sufficient/dependent on others. To be assured that experimental effects could not be taken as an artifact of particular constructs, the first and second groups were randomly alternated as target and non-target constructs. That is, for half of the subjects, the first group became target constructs. For the other half, the second group became target constructs. These constructs were presented on score sheets, with one sheet per career role (see Appendix A). In total, subjects rated 24 career roles on 24 different rating sheets.

Career roles. Twenty-four career roles were selected primarily from the 16 PF Handbook (1970). These career roles were selected to form three groups of eight which are roughly comparable. The first group included: artist, social worker, accountant, electrician, policeman, psychiatric nurse, salesman, and lawyer. The second group included: writer, guidance counsellor, engineer, plumber, airline pilot, primary school teacher, television anchorman for the news, and physician. The third group included: musician, clergyman, banker, carpenter, fireman, special education teacher, newspaper reporter, and business executive. These groups were randomly

assigned to the three phases of the experiment. That is, the first group sometimes appeared on the pre-treatment test, sometimes on the treatment test, and sometimes on the post-treatment test. Similarly, the second and the third groups were randomly placed. Also, the order in which the career roles were presented was random for each subject.

Rating form. As shown in Appendix A, there was one sheet ($8\frac{1}{2}$ x 11 inch white paper) for each role rated. The career role was printed at the top center of the form. The 16 bipolar constructs were listed directly below, with the target constructs listed first, and separated by a line from the eight non-target constructs. Each bipolar dimension was separated by a rating scale. For example, the format approximated the following:

Timid . . . . . Bold

This seven-point scale was placed between the poles of each construct on the rating form. From left to right, the seven points were labelled 'very', 'moderately', 'slightly', 'neither or in between', 'slightly', 'moderately', and 'very'. For example, subjects were instructed to circle the dot which best described the way they construed a given career role. If the fifth dot were circled for instance, it would indicate that the role was seen as slightly bold.

Quantitative feedback. During treatment, subjects received feedback immediately after each career role was rated. Qualitative feedback will be described in the following section. Quantitative feedback was presented by the experi-

menter circling the presumably correct ratings on the subject's rating sheet, using a red marking pencil to contrast more clearly with the subject's ratings. These 'correct' ratings were determined by fixed amounts of divergency which were established for the validation and invalidation conditions. For the validation condition, total rating discrepancies were always ten points per role. For the invalidation condition, rating discrepancies always totaled sixteen points. Discrepancy points were calculated by summing the absolute differences between subject ratings and experimenter ratings over eight target constructs.

To assist the experimenter in providing quantitative feedback, error keys were developed and written on 3 x 5 index cards for use during the experiment. There was one card for the validation condition (Error Card V) and one for the invalidation condition (Error Card I). The cards each contained an 8 x 8 matrix of numbers. These numbers represented discrepancy points between actual ratings on constructs and the presumably correct ratings, i.e. the amount by which they were to be separated. For example, if a subject rated a career role as four on the first construct, and the error key indicated that he should be off by two, then the experimenter's ratings would either have been a two or a six. Direction was not specified. Below are listed the error keys (V and I) for both conditions.

Error Card V Roles	Error Card I Roles	Constructs
1 2 4 2 2 0 0 1	0 4 3 0 4 3 0 4	1
1 2 1 2 1 0 1 1	5 4 3 1 5 3 1 3	2
2 0 0 2 1 3 4 2	4 0 3 3 3 2 2 3	3
0 0 1 1 0 2 2 0	3 2 2 1 0 4 2 2	4
0 0 2 1 0 2 2 2	0 0 2 4 0 0 2 2	5
3 3 0 1 2 2 1 2	0 3 2 3 0 0 3 1	6
2 2 0 1 1 0 0 1	4 3 1 2 4 0 3 1	7
1 1 2 0 3 1 0 1	0 0 0 2 0 4 3 0	8

Note that all columns of error card V sum to ten while all columns of error card I sum to sixteen. The use of these error columns was random. Also, it should be stressed that these error columns are guides rather than fixed rules. For example, suppose a subject rated a role as four on a construct where he was supposed to be off by five. This is impossible. Consequently, the experimenter simply assigned the five to a construct rating where it was possible to do so, and gave that construct the discrepancy score of the one whose place it took.

The error totals of ten and sixteen were developed intuitively. By calling attention to some constructs rather than others, the experimenter both validated and invalidated through qualitative feedback each condition of quantitative feedback. To be able to do this with some improvised conviction, error totals exceeding 16 appeared to be too far off to have anything positive to say about them. Similarly, error totals less than ten appeared to be so accurate that it would be difficult to speak of error, to call attention

to inaccuracy.

Procedure

All subjects were oriented to the experiment with the following remarks:

> "A great number of studies indicate that people in the same occupation tend to have similar personalities. Different personalities are drawn to different types of careers. For example, its easy to see that the average detective is quite different from the average accountant, isn't it? In this study we are interested in seeing how well, how accurately, students perceive the types of persons in different career roles.
>
> Given the fact that there are distinctive personality types in different careers, it is very important to know whether or not students have an accurate view of them. The importance of finding out how accurately students perceive these roles, the type of personality required in different careers, is to try to prevent disastrous career choices. That is, through inaccurate perception, you could choose a career that you are not equipped to handle, which would make you miserable.
>
> For example, it's obvious that different personality characteristics are stressed in different careers. For instance, it might be advantageous for an accountant to be coldly methodical, but not advantageous for a social worker. Or consider a nurse. To be effective, a nurse is often imagined to be warm and sympathetic. But to remain effective, to avoid emotional upset when patients die or fail to improve, a nurse might have to learn to be detached and uninvolved. Because of the complexity of the demands in different careers, the true state of affairs (what these people are like who occupy different roles) is not always obvious.
>
> What I would like you to do is to rate different career roles on this list of personality characteristics. These dimensions of personality have been taken from a personality test (holding up the 16 PF Questionnaire) which has been used to give personality profiles to people in a great number of different careers."

Subjects were then shown the rhetorical aids, with attention informally called to the numerical profiles of different careers and to the personality dimensions. After a brief discussion and informal questions were answered,

instructions continued:

> "Your task is to judge career roles on these dimensions of personality outlined here (holding out the rating sheet). As you will notice, I have given synonyms to some of the dimensions you saw on the test, so that they will be easier to understand. To provide a warm-up and to familiarize you with the task, we'll go through eight different careers, and then I'll begin to give you feedback from the standard profiles to show you how well you are doing."

Subjects were then given the first set of eight rating sheets with career roles printed at the top.

> "The first career is X (indicating the career on the first rating form). What is the average X like? What distinctive characteristics would this person have? Consider the first personality dimension. Circle the point which best indicates what the average X is like in your judgment."

The rating scale was exemplified using the first construct. For example, "this dot means very, this slightly, and so on." Subjects then completed ratings on all 16 constructs, and continued rating the next seven career roles on separate rating sheets.

When the first eight roles were completed, subjects were informed that they would receive feedback on the next eight, and that they could use the feedback to try to improve the accuracy of their judgments. Also, they were informed that feedback would be given only on the first eight construct judgments to save time. Subjects received one of four combinations of feedback.

I: Qualitative validation and quantitative validation

II: Qualitative validation and quantitative invalidation

III: Qualitative invalidation and quantitative validation

IV: Qualitative invalidation and quantitative invalidation

To provide quantitative feedback, the experimenter took the subject's role rating as soon as it was completed, and using either error card V or I, marked the presumably correct answers with a red marking pencil. The subject was given 15 to 30 seconds to study the corrections, and then received qualitative feedback, the experimenter's comments on performance.

Validating qualitative feedback consisted of a variety of experimenter comments which indicated that the subject was fairly accurate in his or her judgments. For example, the experimenter might say: "Your ratings here (calling attention to the most accurate ratings) are very close to the personality profile of this role. You have pinpointed the key traits of people in this occupation, and overall, show an extremely accurate assessment of this career." Another commentary was: "Your judgments are very realistic. You are right on with those judgments (indicating the most accurate ones), and the general trend of judgments is very accurate." In general, commentary called attention to the most accurate judgments (which were either implicitly or explicitly communicated as key characteristics), suggesting that the overall profile or trend of judgments was accurate, and stated that the subject was doing very well.

Invalidating qualitative feedback consisted of experimenter comments which indicated that the subject was rather inaccurate in his or her judgments. For example, the experimenter might say: "In comparison with the actual personality

profile for people in this career, you are quite a bit off on key characteristics (calling attention to the most inaccurate ratings). People are not really X and Y (the most inaccurate judgments), but quite different, as you see." Another commentary was: "What you have been led to believe about people in this career is rather unrealistic. Note the large discrepancies here (pointing to the most inaccurate judgments). The overall trend of judgment is off by a fair amount." In general, commentary called attention to the most inaccurate judgments (suggesting that they were key characteristics), indicated the overall trend of judgments to be inaccurate, and stated explicitly or implicitly that the subject was doing rather poorly.

To assure spontaneity of expression, verbal commentary was not written out precisely, which would either have required rote memorization or reading aloud. Rather, different responses were rehearsed until a variety of responses (different ways of saying the same thing) were established within the experimenter's repertoire, and could be delivered spontaneously with some degree of improvised conviction.

In total, subjects received feedback on eight role ratings in succession, without reference to previous judgments. And the feedback on the first eight constructs was assumed (verbal statement from experimenter to subject) to reflect the way he or she was doing on the remaining eight constructs.

Following treatment, subjects received eight more rating sheets with eight new career roles printed at the top.

"This is the final task. You are to rate eight more career roles. This time, you will be on your own. No feedback will be given. Try to be as accurate as possible. It is very important that students increase their ability to accurately judge the personality types within different careers."

Upon completion of the final ratings, a debriefing session was held. Subjects were informed of the actual purpose of the experiment and told where they could obtain accurate profiles of these career roles. Informal discussion was encouraged concerning the subjects' thoughts and feelings during the experimental session.

Measures

Intensity. For each subject, a separate principal components analysis (BIO MED OIM) was conducted on each pre-treatment, treatment, and post-treatment grid. The size of the first component indicated the greatest amount of variance among ratings that could be accounted for by one dimension. Therefore, the more interrelated the judgments, the larger the first component. The size of the first component of the pre-grid was used to divide subjects into high and low intensity groups, permitting a more detailed investigation of the experimental effects. The percentage of total variance accounted for by the first component was used to measure intensity change within subject judgments. Using similar sized grids, Cochran (Note 1) found a correlation over .98 between the size of the first principal component and Bannister's

measure of intensity. Gain scores (the measure of change used) were obtained by subtracting the percentage of variance score for the target and non-target pre-treatment grids from the corresponding treatment and post-treatment grids respectively. For example, subtracting target pre-treatment grid "percentage of variance" score from the treatment grid "percentage of variance" score yielded the intensity gain score for target construct judgments during treatment.

Consistency. A principal components analysis of the pre-treatment grids also provides orderings of construct loadings on components which correlate zero with one another. These loadings on different components can be thought of as ratings on super-ordinate dimensions. They represent dominant patterns of ratings or judgments found in a grid. Two measures of consistency were derived from these orderings.

First, using principal components regression analysis (BIO MED 02M), the ordering of constructs on the first component was correlated with each column of role judgments on the treatment grid and the post-treatment grid. The ratio of variance accounted for by the first component, under such analysis, to the total amount of variance within a particular column, indicated the extent to which role ratings reflected prior construct organization. The eight percentage scores for the treatment grid were averaged to yield an overall treatment grid consistency measure, and a similar averaging was performed on post-treatment grid scores to give an overall measure of post-treatment consistency. The higher the average,

the more the person is making judgments in accordance with the first principal component of the pre-treatment grid. This scoring procedure was conducted separately for both target and non-target constructs.

Second, using principal-components regression analysis (BIO MED 02M) the ordering of constructs on the first three components was correlated with each column of role judgments on the treatment grid and the post-treatment grid. The statistical procedure employed for the first three components was identical to that followed for the first component analysis. This procedure was also conducted separately for both target and non-target constructs. This second measure permitted consideration of whether differences were due to a more total change of construct relation patterns or to a mere shifting of emphasis from the first component to another dominant dimension of judgment.

On the basis of previous research (Cochran, 1973, 1976, in press) it appeared probable that there might be substantial differences between high and low intensity subjects in the way they modified construct organization. Under invalidation, in fact, high intensity subjects decreased in intensity while low intensity subjects remained unchanged in intensity. Such a pattern can mask significant differences in the above-mentioned analysis. Consequently, all tests were conducted with subjects divided into groups with high and low intensity scores.

CHAPTER III

RESULTS

For reference, experimental groups were numbered in the following manner. Group 1 received qualitative validating and quantitative validating feedback. Group 2 received qualitative validating and quantitative invalidating feedback. Group 3 received qualitative invalidating and quantitative validating feedback. Group 4 received qualitative invalidating and quantitative invalidating feedback.

Pre-treatment Group Characteristics

The pre-treatment Intensity scores for target constructs were: Group 1, $\bar{X}$ = 59.80, S.D. = 13.62; Group 2, $\bar{X}$ = 59.86, S.D. = 13.62; Group 3, $\bar{X}$ = 59.92, S.D. = 12.92, and Group 4, $\bar{X}$ = 58.42, S.D. = 12.28. Using a one-way ANOVA, these group means were not significantly different ($F(3,56)$= 0.04, n.s.).

The pre-treatment Intensity scores for non-target constructs were: Group 1, $\bar{X}$ = 62.06, S.D. = 11.89, Group 2, $\bar{X}$ = 57.97, S.D. = 10.22, Group 3, $\bar{X}$ = 56.93, S.D. = 9.77, and Group 4, $\bar{X}$ = 62.08, S.D. = 12.50. These group means were not significantly different ($F(3,56) = 0.88$, n.s.). Therefore, there appeared to be no significant differences between groups on Intensity scores prior to treatment.

High and Low Intensity Subjects

As expected, the data indicate major differences in the way high and low intensity subjects responded to treatment. Consequently, subjects within each of the four groups were divided into subgroups of high and low intensity subjects, on the basis of pre-treatment target grid intensity scores. The division was made at an Intensity score of 58.5, which was slightly below the median score. This cut-off permitted equal membership in each of the 'high' and 'low' intensity subgroups, for all four experimental subgroups.

Prior to treatment, the experimental groups within both the high and low intensity partitions were quite similar. The pre-treatment target grid scores for high intensity subjects were: Group 1, $\bar{X}$ = 70.36, S.D. = 7.57; Group 2, $\bar{X}$ = 69.17, S.D. = 7.36; Group 3, $\bar{X}$ = 69.08, S.D. = 8.54; Group 4, $\bar{X}$ = 68.49, S.D. = 5.16. Using a one-way ANOVA these subgroup means were not significantly different ($F(3,28) = 0.08$, n.s.). The pre-treatment target grid scores for low intensity subjects were: Group 1, $\bar{X}$ = 47.73, S.D. = 5.71; Group 2, $\bar{X}$ = 49.21, S.D. = 5.29; Group 3, $\bar{X}$ = 48.59, S.D. = 5.15; Group 4, $\bar{X}$ = 46.91, S.D. = 4.78. Again, using a one-way ANOVA there was no significant difference between these subgroups ($F(3,24) = 0.02$, n.s.).

Effects of Varying Validational Evidence

The first eight hypotheses are concerned with the differing effects of validational and invalidational feed-

back upon individual construing processes. Tests of these effects involve comparisons between the first group and the fourth group. Tables 1 and 2 summarize the t-tests conducted for the first eight hypotheses.

Intensity (Hypotheses 1-4). Do validation and invalidation significantly affect the interrelatedness of constructs (i.e. intensity) in a personal construct system? It was predicted that validational and invalidational feedback during treatment would yield significant differences in intensity scores. These differences were hypothesized to persist to a post-treatment phase in the absence of feedback, for target and non-target constructs. Analyses were performed between Groups 1 and 4 on Intensity gain scores. (Intensity gain scores refer to the difference between the amount of variance accounted for by the first component on the pre-treatment grid and the amount of variance accounted for on subsequent grids. For example, subtracting target treatment intensity from target pre-treatment intensity yields the gain score for target change in intensity during treatment.)

Results of t-tests on Intensity gain scores indicate that in general subjects receiving both validating and invalidating feedback decreased the interrelatedness of their constructs. (Note there were two types of t-tests: the first was the standard test between two groups; the second was a test between groups within the analysis of variance using all four groups.) The validated group lost a small amount of

structure while the invalidated group lost radically. This unexpected reaction to feedback of the validated group will be further investigated by an extended analysis to be presented later. For the present, t-scores (see Table 1) indicate that for high intensity subjects, validation and invalidation produced significantly different effects on the intensity of target constructs during and following treatment. Also, with non-target constructs, validation and invalidation resulted in significantly different intensity scores following treatment. During treatment, the difference approached significance and following treatment, reached significance. The t-scores for Intensity gain score differences between the validated group and the invalidated group of low intensity subjects show no significant differences (see Table 2). Generally then, significant differences occurred between validated and invalidated high intensity subjects, but not between similar groups of low intensity subjects.

Consistency (Hypotheses 5-8). Do validation and invalidation significantly affect the consistency of construct relations in a personal construct system? It was predicted that significant differences would be found in consistency scores by varying validational feedback (treatment phase) and would be maintained through a post-treatment phase. These predictions were made for both target and non-target constructs. Statistical analyses for consistency measures were performed for both the first component (C_1) and the first three components (C_{1-3}) levels. The first measure traces the consistency

TABLE 1

Hypotheses 1-8: Means, standard deviations, and t-ratios for high intensity subjects

Hypotheses and Measures	Validated Group Mean	Validated Group S.D.	Invalidated Group Mean	Invalidated Group S.D.	t_1	t_2
1) Target Intensity Change: Treatment	-9.71	9.76	-26.59	7.60	3.86**	2.77**
2) Non-target Intensity Change: Treatment	-1.51	10.06	-11.34	10.50	1.91	1.45
3) Target Intensity Change: Post-treatment	-3.85	12.40	-25.32	9.81	3.84**	2.74**
4) Non-target Intensity Change: Post-treatment	0.86	7.53	-16.79	12.57	3.41**	2.98**
5a) Target Consistency: Treatment (C_1)	53.58	13.18	26.85	6.74	5.11**	4.16**
5b) Target Consistency: Treatment (C_{1-3})	72.90	8.73	60.88	12.03	2.29*	2.00
6a) Non-target Consistency: Treatment (C_1)	62.18	14.32	36.25	22.00	2.79**	3.36**
6b) Non-target Consistency: Treatment (C_{1-3})	80.54	10.43	59.99	19.48	2.63*	3.19**
7a) Target Consistency: Post-treatment (C_1)	55.19	21.80	25.67	12.94	3.29**	3.09**
7b) Target Consistency: Post-treatment (C_{1-3})	73.45	14.95	59.77	14.33	2.55*	1.85
8a) Non-target Consistency: Post-treatment (C_1)	57.79	18.34	34.77	14.49	2.79**	3.08**
8b) Non-target Consistency: Post-treatment (C_{1-3})	77.90	11.75	59.25	12.38	3.09**	3.77**

Note. t_1 refers to standard t-test between groups (df = 14) while t_2 refers to t-test using the error term from the analysis of variance (df = 28). C_1 refers to consistency score from first component analysis while C_{1-3} refers to consistency score from the first three components analysis.

*$p < .05$

**$p < .01$

TABLE 2

Hypotheses 1-8: Means, standard deviations, and t-ratios for low intensity subjects

Hypotheses and Measures	Validated Group Mean	Validated Group S.D.	Invalidated Group Mean	Invalidated Group S.D.	t_1	t_2
1) Target Intensity Change: Treatment	-3.64	5.05	-4.91	4.18	0.51	0.32
2) Non-target Intensity Change: Treatment	-7.20	9.40	-0.60	13.93	-1.03	-1.36
3) Target Intensity Change: Post-treatment	2.54	10.70	1.95	9.02	0.09	0.08
4) Non-target Intensity Change: Post-treatment	-1.66	14.78	-6.20	15.35	0.56	0.58
5a) Target Consistency: Treatment (C_1)	25.56	15.36	29.57	8.97	-0.60	-0.56
5b) Target Consistency: Treatment (C_{1-3})	65.93	11.60	65.55	10.57	0.06	0.07
6a) Non-target Consistency: Treatment (C_1)	36.28	12.14	42.59	12.95	-0.94	-0.84
6b) Non-target Consistency: Treatment (C_{1-3})	69.29	7.92	67.70	11.98	0.29	0.33
7a) Target Consistency: Post-treatment (C_1)	31.72	14.58	24.97	12.41	0.93	0.84
7b) Target Consistency: Post-treatment (C_{1-3})	65.71	12.79	59.75	9.75	0.98	0.98
8a) Non-target Consistency: Post-treatment (C_1)	36.49	16.72	41.26	18.70	-0.50	-0.55
8b) Non-target Consistency: Post-treatment (C_{1-3})	66.34	9.63	63.86	11.74	0.43	0.36

Note. t_1 refers to standard t-test between groups (df = 12) while t_2 refers to t-test using the error term from the analysis of variance (df = 24). C_1 refers to consistency score from the first component analysis while C_{1-3} refers to consistency score from the first three components analysis.

of judgments using the first component or major dimension of variation as a standard. The second measure allows consideration of another question. That is, are differences occurring in Consistency due to a more total change of construct relation patterns or are they due to shifting of emphasis from the first component to other major dimensions of judgment? (The total of the first three components accounts for an average of 90.8 percent of the variance in this study.)

The t-scores obtained from a comparison of C_1 scores for high intensity subjects indicate significant differences between validational and invalidational groups (i.e. Groups 1 and 4). These differences occurred on both treatment and post-treatment measures for target and non-target constructs (see Table 1).

Results of a comparison of C_{1-3} scores for high intensity subjects (t-tests) also showed significant differences between Groups 1 and 4. Again significant differences occurred for target and non-target constructs both during and following treatment. It should be noted that on two occasions the significant effect was not confirmed by ANOVA t-tests (see Table 1). Since all four groups were involved in ANOVA tests this effect was produced by the larger error term. The variance of the mixed feedback groups was much larger than that of the other groups, with some of the mixed-feedback groups' subjects responding like the validated subjects, and others responding as did the invalidated subjects. For low intensity subjects, t-tests for Consistency score differences

between the validated and invalidated groups did not reach significance (see Table 2). In summary, significant differences occurred between validated and invalidated high intensity subjects on Consistency measures, but not between similar groups of low intensity subjects.

Feedback type. Hypotheses 9 and 10 predicted that the two types of feedback (qualitative and quantitative) would contribute independently to the effects of validation and invalidation upon Intensity and Consistency measures. A 2 x 2 ANOVA (qualitative x quantitative) indicated that main effects of feedback type contributed significantly to Intensity and Consistency scores, but there appears to be no trend relevant to the pattern of effects (i.e. the significant effects varied inconsistently between experimental phases and between target and non-target construct judgments). There were no significant interaction effects (see Tables 3 and 4). In addition, no significant differences were found between groups 2 and 3 on measures of Intensity and Consistency (for either high or low intensity subjects). The lack of significance here where the groups received a combination of the different forms of validational and invalidational feedback (see Tables 5 & 6), suggests that the mixed feedback effect was approximately similar across feedback combinations. This, considered with the lack of interaction and the presence of main effects, provides tentative support for an additive as opposed to an interactive feedback effect.

TABLE 3

Qualitative vs quantitative feedback: A summary of 2-way ANOVAs with high intensity subjects

Measures	Main Effects: Qualitative	Main Effects: Quantitative	Interaction
1) Target Intensity Change: Treatment	5.03*	2.79	0.65
2) Non-target Intensity Change: Treatment	0.95	1.90	0.74
3) Target Intensity Change: Post-treatment	3.06	4.49*	0.02
4) Non-target Intensity Change: Post-treatment	0.51	12.25**	2.78
5a) Target Consistency: Treatment (C_1)	11.38**	6.28*	1.39
5b) Target Consistency: Treatment (C_{1-3})	3.86	0.76	0.06
6a) Non-target Consistency: Treatment (C_1)	5.60*	5.73*	0.08
6b) Non-target Consistency: Treatment (C_{1-3})	2.61	8.39**	1.31
7a) Target Consistency: Post-treatment (C_1)	6.28*	3.47	0.39
7b) Target Consistency: Post-treatment (C_{1-3})	1.23	2.27	0.14
8a) Non-target Consistency: Post-treatment (C_1)	1.57	9.63**	0.18
8b) Non-target Consistency: Post-treatment (C_{1-3})	8.16**	6.08*	0.94

Note. C_1 refers to consistency score from the first component analysis while C_{1-3} refers to consistency score from the first three components analysis. The degrees of freedom for all tests was 1,28.

*$p < .05$
**$p < .01$

TABLE 4

Qualitative vs quantitative feedback: A summary of 2-way ANOVAs with low intensity subjects

Measures	Main Effects		Interaction
	Qualitative	Quantitative	
1) Target Intensity Change: Treatment	0.37	1.12	0.58
2) Non-target Intensity Change: Treatment	0.05	2.93	1.41
3) Target Intensity Change: Post-treatment	0.06	0.12	0.13
4) Non-target Intensity Change: Post-treatment	0.84	0.01	0.00
5a) Target Consistency: Treatment (C_1)	0.05	0.32	0.94
5b) Target Consistency: Treatment (C_{1-3})	0.03	0.08	0.20
6a) Non-target Consistency: Treatment (C_1)	0.01	1.14	0.04
6b) Non-target Consistency: Treatment (C_{1-3})	0.01	0.14	0.52
7a) Target Consistency: Post-treatment (C_1)	1.92	0.04	0.48
7b) Target Consistency: Post-treatment (C_{1-3})	1.71	0.01	1.89
8a) Non-target Consistency: Post-treatment (C_1)	0.20	1.50	0.63
8b) Non-target Consistency: Post-treatment (C_{1-3})	0.33	0.00	0.03

Note. C_1 refers to consistency score from first component analysis while C_{1-3} refers to consistency score from first three components analysis. The degrees of freedom for all tests was 1,24.

TABLE 5

Means, standard deviations, and t-ratios for high intensity subjects in mixed feedback treatment groups

Measures	Group 2 Mean	Group 2 S.D.	Group 3 Mean	Group 3 S.D.	t_1	t_2
1) Target Intensity Change: Treatment	-15.90	15.07	-13.43	14.68	-0.33	-0.40
2) Non-target Intensity Change: Treatment	- 1.98	14.52	- 3.67	12.28	0.25	0.25
3) Target Intensity Change: Post-treatment	-14.24	18.11	-16.30	20.21	-0.21	-0.26
4) Non-target Intensity Change: Post-treatment	4.86	11.50	- 6.81	14.54	1.77	1.97
5a) Target Consistency: Treatment (C_1)	43.60	14.39	47.50	15.34	-0.53	-0.61
5b) Target Consistency: Treatment (C_{1-3})	65.65	15.05	70.23	11.81	-0.67	-0.76
6a) Non-target Consistency: Treatment (C_1)	47.77	9.45	47.62	14.35	0.02	0.02
6b) Non-target Consistency: Treatment (C_{1-3})	78.39	6.89	72.56	11.37	1.24	0.91
7a) Target Consistency: Post-treatment (C_1)	34.05	17.07	38.39	22.97	-0.43	-0.45
7b) Target Consistency: Post-treatment (C_{1-3})	62.84	12.52	59.99	11.89	-0.56	-0.28
8a) Non-target Consistency: Post-treatment (C_1)	53.58	14.27	43.61	11.98	1.48	1.31
8b) Non-target Consistency: Post-treatment (C_{1-3})	71.29	7.50	72.66	6.73	-0.38	-0.28

Note. Group 2 received qualitative validating/quantitative invalidating feedback while Group 3 received qualitative invalidating/quantitative validating feedback. t_1 refers to standard t-test between groups (df = 14) while t_2 refers to t-test using the error term from the analysis of variance (df = 28). C_1 refers to consistency score from the first component analysis while C_{1-3} refers to consistency score from the first three components analysis.

TABLE 6

Means, standard deviations, and t-ratios for low intensity subjects in mixed feedback treatment groups

Measures	Group 2 Mean	Group 2 S.D.	Group 3 Mean	Group 3 S.D.	t_1	t_2
1) Target Intensity Change: Treatment	-0.21	9.09	-4.47	9.81	0.93	1.18
2) Non-target Intensity Change: Treatment	-2.39	10.26	2.72	4.83	-1.19	-1.06
3) Target Intensity Change: Post-treatment	1.94	11.95	-0.72	14.84	0.37	0.42
4) Non-target Intensity Change: Post-treatment	-6.89	10.88	-1.30	16.86	-0.74	-0.71
5a) Target Consistency: Treatment (C_1)	31.64	13.76	33.32	14.11	-0.22	-0.23
5b) Target Consistency: Treatment (C_{1-3})	64.93	9.90	63.28	6.06	0.38	0.32
6a) Non-target Consistency: Treatment (C_1)	35.85	13.46	40.87	17.20	-0.61	-0.67
6b) Non-target Consistency: Treatment (C_{1-3})	71.46	9.44	70.48	5.56	0.24	0.20
7a) Target Consistency: Post-treatment (C_1)	27.79	14.43	36.79	18.20	-1.03	-1.12
7b) Target Consistency: Post-treatment (C_{1-3})	65.98	12.92	71.31	9.59	-0.88	-0.88
8a) Non-target Consistency: Post-treatment (C_1)	28.92	8.13	39.13	19.14	-1.31	-1.29
8b) Non-target Consistency: Post-treatment (C_{1-3})	64.74	14.74	67.53	14.84	-0.39	-0.48

Note. Group 2 received qualitative validating/quantitative invalidating feedback while Group 3 received qualitative invalidating/quantitative validating feedback. t_1 refers to standard t-test between groups (df = 12) while t_2 refers to t-test using the error term from the analysis of variance (df = 24). C_1 refers to consistency score from the first component analysis while C_{1-3} refers to consistency score from the first three components analysis.

However, it should be noted that Intensity scores for the mixed feedback treatment groups were not accurately reflected in the means. That is, the scores of the individual subjects in both of the mixed feedback groups were typical of either validated or invalidated individual subjects' scores, not of an average between both sets. But, consistency scores for the mixed feedback groups were generally typical of the average score. Consequently, at least for intensity scores, subjects appeared to accept one type of feedback or the other rather than both, which discredits an additive interpretation.

Extended analysis. Although the analyses reported above support strongly the hypotheses concerning the differential effects of validation and invalidation, an extended analysis (using correlated t-tests) was conducted to assess changes in patterns of construing for all groups regardless of level of intensity. Table 7, which reports the results of correlated t-tests conducted on Intensity data, shows comparisons between pre- and post-treatment data, as well as between pre-treatment and treatment, and between treatment and post-treatment, for both target and non-target constructs. These results indicate the extent and direction of change in personal construct system organization, and an effect of the varying validational treatments. For all groups the mean loss in group-Intensity scores was significant from the pre-treatment to the treatment phase on target constructs. The only case where this loss was maintained through post-

TABLE 7

Intensity: Means, standard deviations, and correlated t-tests of change for each group separately

Grids	Pre-treatment Mean	Pre-treatment S.D.	Treatment Mean	Treatment S.D.	Post-treatment Mean	Post-treatment S.D.	t-ratios t_1	t_2	t_3
Group 1: Target Construct Judgments	59.80	13.62	52.92	13.72	58.89	14.21	3.22**	-2.74*	0.30
Non-target Construct Judgments	62.06	11.89	57.89	14.22	61.74	13.87	1.62	-1.11	0.13
Group 2: Target Construct Judgments	59.86	13.62	50.85	13.97	50.83	15.48	2.47*	-0.06	1.93
Non-target Construct Judgments	57.97	10.22	57.28	10.78	54.40	9.71	0.27	0.94	0.94
Group 3: Target Construct Judgments	59.92	12.92	51.14	8.29	52.83	12.34	2.21*	-0.47	1.51
Non-target Construct Judgments	56.93	9.77	54.75	9.81	56.30	11.21	0.69	-0.50	0.19
Group 4: Target Construct Judgments	58.42	12.28	41.95	4.14	45.83	7.27	5.02**	-1.85	3.15**
Non-target Construct Judgments	62.08	12.50	55.76	11.92	50.23	10.62	1.89	1.43	3.17**

Note. t_1 refers to change in intensity (df = 14) from pre-treatment to treatment conditions. t_2 refers to change in intensity (df = 14) from treatment to post-treatment conditions. t_3 refers to change in intensity (df = 14) from pre-treatment to post-treatment conditions.

*$p < .05$
**$p < .01$

treatment was for the invalidated group. For the invalidated group only, a significant mean Intensity score loss occurred also for non-target constructs. The trend of loss of structure from pre-treatment to treatment, though not significant, was continued from treatment to post-treatment. Though this latter comparison was not significant, the pre- to post-treatment intensity loss was significant. Treatment to post-treatment mean changes in Intensity reached significance only for the validated group (Group 1). Here the significant loss of structure from pre-treatment to treatment was regained (see Table 7). In summary then, there appeared to be significant group 'mean' changes throughout only for the invalidated group (Group 4), where the loss of structure occurred for both target and non-target construct judgments.

CHAPTER IV

DISCUSSION

The only significant effects of invalidation demonstrated in this study involved high intensity subjects. Both construct intensity and consistency were significantly decreased by invalidational evidence. This loosening of construct organization was apparent not simply with the constructs and roles which were invalidated, but with constructs and roles which were not invalidated. That is, loosening was pervasive rather than isolated.

Serial invalidation of high intensity subjects results in predictions or judgments which were more random and which failed to make substantial use of prior construct organization. Behavioral observations during treatment were supportive of this loosening process. Subjects took more time in making judgments during treatment, frequently erased ratings, and moved their pencils from one rating point to another as if each position on the scale was being carefully evaluated.

Some comments from invalidated subjects are also revealing of the treatment impact. During treatment, one subject commented that "there's no point; I don't know what's happening to me.",while another exclaimed "How can I be that wrong? I'm hopeless.". Following the debriefing session, one subject commented that "you really had me wondering.",

while another remarked that "I couldn't think straight after a while."

These observations and comments suggest a developing uncertainty and bewilderment among invalidated subjects, that complements the disorganized use of personal constructs which the measures of intensity and consistency document. The comments might indicate their subjective realization of the diminishing reliability and usefulness of their construct systems.

As an analogue study concerned with the generation of thought disorder, the evidence demonstrates that invalidation exerted a strong, disruptive influence upon persons' construct systems. This effect is certainly similar in direction to clinical descriptions of thought disorder. Constructs were applied in a less consistent and less coordinated fashion and invalidated subjects appeared to be uncertain about how to respond. However, parallel studies which are more naturalistic will be necessary to connect this evidence firmly with the development of thought disorder.

One interesting question which arises from this study is why high intensity subjects were so strongly affected by invalidation while low intensity subjects did not appear to be affected at all (at least on the measures used). Since low intensity subjects cannot afford to lose further structure and maintain a workable construct system (c.f. Cochran, in press), they might discount invalidational evidence more effectively than high intensity subjects. Or, maybe they

expect a higher degree of invalidation than tight construers and consequently are not as disrupted when it occurs.

Assuming that both high and low intensity subjects do accept the feedback that their judgments are erroneous, one possible explanation is that disruptiveness is a function of construct interrelatedness. For example, research by Levy (1956), Hinkle (1965), and more recently Crockett and Meisel (1974) assessed the relationship between construct interrelatedness and the degree of change made in interpretation of events. In general, these studies demonstrated that even when change appeared desirable, highly interconnected constructs or construct systems were resistent to change. But when the change did occur, it only happened with massive alterations in the construct system. As a consequence of high degrees of interrelatedness, then, invalidation of even a single construct is more apt to ramify throughout the system of constructs (Weick, 1968). However, for subjects with low intensity, i.e. subjects with weakly interrelated constructs, invalidation might be more manageable since the implications of change per construct are less pervasive.

Cochran's (1976) study of similarity/difference orientations might also be relevant. In his study, subjects who perceived more similarities among contradictory 'stimulus persons' maintained or gained structure while subjects who perceived more differences lost structure. For individuals who tend toward interpretation based on similarities, invalidation on one channel of construing might be more likely to

influence the construing pattern of a large number of other judgments than it would for individuals who emphasize differences.

This striking difference between high and low intensity subjects might also be of value in future therapeutic strategies. For example, while high intensity subjects might be studied to determine how invalidation disrupts construing, low intensity subjects might be studied to determine how it might be managed more effectively.

A second, unexpected question which arises from this study is why validated subjects decreased in intensity from the pre-grid to the treatment grid. Since they returned to a customary degree of intensity on the post-grid, it seems clear that the validation treatment condition produced only a transient loss of structure.

One explanation might be that treatment was similar to a testing situation and that any testing situation in which something is perceived to be at stake will decrease construct intensity. For instance, subjects might stress accuracy even at the expense of coherence. Another explanation might be that the validating treatment condition was, if fact, mildly invalidating. A ten-point error-rate per role could have been too great to be considered validating, in spite of the positive experimenter comments. Also, since the error-rate was constant, the fact that they did not improve over roles could have been mildly invalidating as well. Occasionally, subjects mentioned that despite their

attempts to use the feedback presented to develop more accurate judgments, they were never able to rate a career role with absolute precision. Both possibilities are plausible and testable. They also serve as a caution in future experiments as the loss of structure of presumably validated subjects stresses the fact that subjects are apt to vary considerably in what they consider to be validating and invalidating.

A third issue concerns the difficulty in interpreting the mixed feedback groups. Statistically, qualitative and quantitative feedback contributed additively (if somewhat inconsistently) to the changes in construct organization of high intensity subjects. On measures of both intensity and consistency, the average for the mixed feedback groups always fell between the averages for the validated and invalidated groups. Also, no interaction approached significance and no t-test between the mixed feedback groups was significant. However, the averages for the mixed feedback groups on measures of intensity were unrepresentative of the actual scores. Subjects tended to either respond similar to validated subjects (minimal loss) or similar to invalidated subjects (huge loss). (See Appendix B for a listing of actual scores for each group under each condition.) But the averages for the mixed feedback groups on measures of consistency were representative of the actual scores, which did tend to fall between the actual scores of validated and invalidated subjects.

Behavioral observations of subjects in the mixed feedback groups support the differential response of intensity scores. For example, some subjects responded slowly as did the invalidated subjects (discussed previously) while others responded with obvious confidence. One subject commented that "I'm really doing much better than I had expected.", while another mentioned that "I knew I'd do well on that one - my brother's a fireman.". Why intensity scores were polarized and consistency scores were more uniform is simply a puzzle at this time. However, the fact that there was a differential response suggests that the strong effect upon invalidated subjects did not result so much from 'double-barrelled' invalidation as from an acceptance of one type of feedback or the other. Perhaps some subjects are sensitive to evaluative comments from people in authority while others are sensitive to actual objective evidence. In this case, qualitative and quantitative feedback are not so much additive as 'either/or' (either the subject accepts one or the other with the same effect upon construct organization). As Bannister (Note 2) suggested, the two types of feedback might have impact upon the credibility of validation to the subject rather than upon disruptive potential of the information. For future studies, the implications are clear. Since subjects are apt to vary in the types of feedback which they will accept, both quantitative and qualitative feedback should be included within experiment designs.

In conclusion, the evidence clearly demonstrates that invalidation loosened the construct systems of high intensity subjects. The power of this demonstration can be plausibly attributed to realism of the experimental tasks, the credibility of the feedback, and the personal relevance of the experiment (i.e. subjects were involved; it mattered to them whether or not they were accurate in judging career roles). Behavioral and construing patterns displayed by invalidated subjects in this study clearly suggest that investigation of natural-environment invalidational processes would contribute toward a further understanding of the construct system disruption that results in thought-disorder. In demonstrating that loosening results from invalidation, this study also provides a much-needed replication of Bannister's (1965) fourth serial invalidation experiment and supplies an informed basis for extensions to naturalistic situations.

REFERENCE NOTES

1. Cochran, L. Personal communication, April 1977.

2. Bannister, D. Personal communication, February 1977.

REFERENCES

Annett, J. Feedback and Human Behavior. Baltimore, Maryland: Penguin Books Inc., 1969.

Bandura, A. Behavior theory and the models of man. American Psychologist, 1974, 29, 12, 859-869.

Bannister, D. Conceptual structure in thought-disordered schizophrenics. Journal of Mental Science, 1960, 106, 1230-1249.

Bannister, D. The nature and measurement of schizophrenic thought disorder. Journal of Mental Science, 1962, 108, 825-842.

Bannister, D. The genesis of schizophrenic thought disorder: A serial invalidation hypothesis. British Journal of Psychiatry, 1963, 109, 680-686.

Bannister, D. The genesis of schizophrenic thought disorder: A re-test of the serial invalidation hypothesis. British Journal of Psychiatry, 1965, 111, 377-382.

Bannister, D. & Mair, J. The Evaluation of Personal Constructs. New York: Academic Press, 1968.

Bilodeau, E.A. & Bilodeau, I.McD. Variable frequency of knowledge of results and the learning of a simple skill. Journal of Experimental Psychology, 1958, 55, 379-383.

Bolles, R.C. Reinforcement, expectancy and learning. Psychological Review, 1972, 79, 5, 394-409.

Breland, K. & Breland, M. The misbehavior of animals. American Psychologist, 1961, 16, 681-684.

Brown, P.L. & Jenkins, H.M. Auto-shaping of the pigeon's key-peck. Journal of the Experimental Analysis of Behavior, 1968, 11, 1-8.

Bruner, J.S., Goodnow, J.J. & Austin, G.A. A Study of Thinking. New York: John Wiley & Sons, 1956.

Cattell, R.B. Personality pinned down. Psychology Today, 1973 (July), 40-46.

Cattell, R.M., Eber, H.W. & Taksuoka, M.M. Handbook for the Sixteen Personality Factor Questionnaire. Chicago: Institute for Personality and Ability Testing, 1970.

Cochran, L. The effects of integrating structurally incompatible information on cognitive organization. (Unpublished doctoral dissertation, University of Utah, 1973). Dissertation Abstracts International, 1973, 34, 2923B-2924B. (University Microfilms No. 73-29,390).

Cochran, L. Categorization and change in conceptual relations. Canadian Journal of Behavioral Science, 1976, 30, 275-286.

Cochran, L. Inconsistency and change in conceptual organization. British Journal of Medical Psychology, in press.

Crockett, W.H. & Meisel, P. Construct connectedness, strength of disconfirmation, and impression change. Journal of Personality, 1974, 42, 2, 290-299.

Estes, W.K. Reinforcement in human learning. In J.T. Tapp, Reinforcement and Behavior. New York: Academic Press, 1969, 63-94.

Frankel, J., Levine, M. & Karpf, D. Human discrimination learning: A test of the blank-trials assumption. Journal of Experimental Psychology, 1970, 83, 342-348.

Gholson, B., Levine, M. & Phillips, S. Hypotheses, strategies and stereotypes in discrimination learning. Journal of Experimental Child Psychology, 1972, 13, 423-446.

Hamilton, G.V. A study of trial and error reactions in mammals. Journal of Animal Behavior Monographs, 1911, 1, 33-66.

Hinkle, D.N. The change of personal constructs from the viewpoint of a theory of construct implications. (Unpublished doctoral dissertation, Ohio State University, 1965). Dissertation Abstracts, 1966, 26, 6837. (University Microfilms No. 66-1790).

Kelly, G.A. The Psychology of Personel Constructs, Vols. 1 & 2. New York: W.W. Norton & Company, Inc., 1955.

Kirk, R.E. Experimental Design: Procedures for the Behavioral Sciences. California: Brooks/Cole Publishing Company, 1968.

Krechevsky, I. "Hypothesis versus chance" in the presolution period in sensory discrimination learning. University of California Publications in Psychology, 1932, 6, 27-44. (b)

Lashley, K.S. Brain Mechanisms and Intelligence. Chicago: University of Chicago Press, 1929.

Levine, M. Mediating processes in humans at the onset of discrimination learning. Psychological Review, 1963, 70, 254-276.

Levine, M. Hypothesis behavior by humans during discrimination learning. Journal of Experimental Psychology, 1966, 71, 3, 331-338.

Levine, M. Human discrimination learning: the subset sampling assumption. Psychological Bulletin, 1970, 74, 397-404.

Levine, M. A Cognitive Theory of Learning: Research on Hypothesis Testing. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975.

Levine, M., Leitenberg, H., & Richter, M. The blank-trials law: The equivalence of positive reinforcement and non-reinforcement. Psychological Review, 1964, 71, 94-103.

Levy, L.H. Personal constructs and predictive behavior. Journal of Abnormal and Social Psychology, 1956, 53, 54-58.

Lincoln, R.S. Learning a rate of movement. Journal of Experimental Psychology, 1954, 47, 465-470.

Radley, A.R. Schizophrenic thought disorder and the nature of personal constructs. British Journal of Social and Clinical Psychology, 1974, 13, 315-327.

Restle, F. A note on the "hypothesis" theory of discrimination learning. Psychological Reports, 1960, 7, 194.

Restle, F. Statistical methods for a theory of cue learning. Psychometrika, 1961, 26, 291-306.

Restle, F. The selection of strategies in cue learning. Psychological Review, 1962, 69, 329-343.

Smode, A. Learning and performance in a tracking task under two levels of achievement information feedback. Journal of Experimental Psychology, 1958, 56, 297-304.

Sokolov, E.N. The modeling properties of the nervous system. In M. Cole and I. Maltzman (Eds.), A Handbook of Contemporary Soviet Psychology. New York: Basic Books, Inc., 1969.

Tolman, E.C. There is more than one kind of learning. Psychological Review, 1949, 56, 144-155.

Trowbridge, M.H. & Cason, H. An experimental study of Thorndike's 'Theory of Learning'. Journal of General Psychology, 1932, 7, 245-258.

Weick, K.E. Processes of ramification among cognitive links. In R.P. Abelson, E. Aronson, W.J. McGuire, T.N. Newcomb, M.J. Rosenberg, and P.J. Tannenbaum (Eds.), Theories of Cognitive Consistency. Chicago: Rank McNally, 1968.

Williams, D.R. & Williams, H. Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. Journal of the Experimental Analysis of Behavior, 1969, 12, 511-520.

Yerkes, R.M. The mental life of monkeys and apes: A study of the ideational behavior in man and other animals. Proceedings of the National Academy of Sciences, 1916, 631-633. Cited by K.W. Spence, An experimental test of the continuity and non-continuity theories of discrimination learning. Journal of Experimental Psychology, 1945, 35, 253-266.

APPENDIX A

CAREER ROLE: _____

	very	moderately	slightly	neither	slightly	moderately	very	
Slack	.	.	.	.	.	.	.	Conscientious
Concrete Reasoning	.	.	.	.	.	.	.	Abstract Reasoning
Timid	.	.	.	.	.	.	.	Bold
Naive	.	.	.	.	.	.	.	Shrewd
Tough-minded	.	.	.	.	.	.	.	Tender-minded
Relaxed	.	.	.	.	.	.	.	Tense
Undisciplined, Acts on Impulse	.	.	.	.	.	.	.	Controlled, Thinks Before Acting
Self-sufficient	.	.	.	.	.	.	.	Dependent on Others
Reserved	.	.	.	.	.	.	.	Warm-hearted
Serious	.	.	.	.	.	.	.	Happy-go-lucky
Easily Led	.	.	.	.	.	.	.	Assertive
Conservative	.	.	.	.	.	.	.	Experimental
Trusting	.	.	.	.	.	.	.	Suspicious
Practical	.	.	.	.	.	.	.	Imaginative
Self-assured	.	.	.	.	.	.	.	Apprehensive
Emotionally Unstable	.	.	.	.	.	.	.	Emotionally Stable

APPENDIX B

TARGET CONSTRUCT 'TREATMENT' CHANGE IN INTENSITY SCORES

High Intensity Subjects:

Group I: -11.41, 2.96, -6.89, -0.01, -9.50, -21.27, -5.95, -25.64
Group II: 0.55, -29.27, 1.86, -26.65, 5.58, -9.22, -24.94, -25.32
Group III: -34.90, -32.49, -25.76, -20.23, -1.35, -9.74, -10.41, 7.71
Group IV: -38.55, -15.10, -26.05, -20.62, -28.95, -24.10, -24.03, -35.28

Low Intensity Subjects:

Group I: -8.42, 0.20, 5.74, -3.75, -6.61, -7.77, -4.85
Group II: -3.33, 8.36, -12.59, -9.10, 7.53, -17.89, -4.28
Group III: 4.03, 6.21, 12.84, -9.67, -7.37, -9.99, 5.41
Group IV: -7.65, -6.04, 0.49, -11.41, -0.06, -4.07, -5.64

TARGET CONSTRUCT 'POST-TREATMENT' CHANGE IN INTENSITY SCORES

High Intensity Subjects:

Group I: 3.66, -4.05, 21.45, 0.20, -12.59, -11.58, -17.41, -10.47
Group II: 0.17, -19.28, -42.29, -23.78, -2.62, 1.55, 3.19, -47.31
Group III: -36.35, -33.18, -19.81, -6.40, -18.81, -21.38, 9.05, 12.97
Group IV: -36.45, -16.44, -24.25, -6.78, -29.35, -27.83, -25.99, -35.50

Low Intensity Subjects:

Group I: 1.09, 14.66, 11.16, 12.84, -12.18, -8.49, -1.91
Group II: -6.78, 7.73, -13.86, 3.92, 25.64, -18.59, -3.11
Group III: 9.40, -10.94, -0.67, 9.49, 20.38, -12.95, -1.14
Group IV: 18.42, -5.46, 10.51, 0.16, -2.55, -4.62, -2.78

NON-TARGET CONSTRUCT 'TREATMENT' CHANGE IN INTENSITY SCORES

High Intensity Subjects:

Group I: -18.60, -3.27, 3.45, 5.91, -5.53, -11.33, 11.38, 5.89
Group II: 5.45, -13.14, 4.32, -23.09, -1.79, 13.70, -0.07, -14.81
Group III: 16.13, -19.30, -3.64, -9.45, -15.02, 5.02, 20.48, -10.09
Group IV: -12.64, 2.65, -21.34, -9.20, -13.19, 0.49, -8.29, -29.13

Low Intensity Subjects:

Group I: -6.02, -1.06, -11.06, -6.62, -26.97, -0.28, 1.03
Group II: 10.29, -0.91, -3.18, 2.40, 8.05, 0.38, 2.00
Group III: -11.13, -20.87, 4.29, 3.50, 6.65, -3.85, 4.68
Group IV: 6.40, 5.63, 4.82, -23.17, 14.99, 4.17, -17.07

NON-TARGET CONSTRUCT 'POST-TREATMENT' CHANGE IN INTENSITY SCORES

High Intensity Subjects:

Group I: -7.24, 10.51, -6.89, -0.51, 10.12, -4.88, -2.09, 7.85
Group II: -3.90, -19.15, 15.37, -22.31, -0.98, 4.64, -1.11, -27.07
Group III: 12.18, -18.14, -6.80, 6.66, 6.73, 15.51, 10.18, 12.57
Group IV: -30.78, -23.58, -11.86, -19.53, 7.94, -22.51, -7.29, -26.71

Low Intensity Subjects:

Group I: -3.05, 4.09, -12.94, 20.55, -7.36, -23.44, 10.53
Group II: -3.92, 5.54, -15.09, 12.07, 24.23, -26.00, -5.92
Group III: -12.50, -18.58, -5.74, -13.16, 2.97, -13.36, 12.16
Group IV: 2.46, -7.55, 11.01, -16.69, -15.44, 11.85, -29.01

TARGET CONSTRUCT 'TREATMENT' CONSISTENCY (C_1) SCORES

High Intensity Subjects:

Group I: 59.49, 49.57, 57.00, 70.86, 61.16, 30.37, 60.93, 39.26
Group II: 56.38, 44.13, 61.46, 27.06, 71.84, 48.54, 29.92, 41.01
Group III: 59.78, 57.27, 40.96, 18.85, 29.58, 50.84, 53.58, 37.91
Group IV: 13.53, 23.43, 34.46, 27.52, 26.21, 30.93, 33.89, 24.82

Low Intensity Subjects:

Group I: 25.09, 32.86, 50.15, 19.95, 34.77, 4.41, 11.66
Group II: 26.68, 37.02, 22.87, 22.10, 63.84, 25.53, 35.23
Group III: 18.71, 44.30, 35.51, 50.73, 11.02, 31.52, 29.68
Group IV: 23.97, 35.73, 36.73, 37.76, 30.93, 29.39, 12.46

TARGET CONSTRUCT 'POST-TREATMENT' CONSISTENCY (C_1) SCORES

High Intensity Subjects:

Group I: 68.72, 73.96, 73.47, 78.52, 46.75, 17.45, 37.31, 45.36
Group II: 34.49, 23.10, 65.15, 32.75, 64.53, 62.89, 8.04, 16.18
Group III: 63.29, 30.36, 50.81, 16.69, 10.71, 30.34, 31.41, 38.78
Group IV: 12.66, 36.68, 36.10, 44.74, 28.30, 22.53, 16.05, 8.33

Low Intensity Subjects:

Group I: 28.77, 40.16, 57.76, 26.97, 16.34, 35.87, 16.17
Group II: 23.09, 21.04, 32.95, 32.86, 74.90, 30.49, 42.18
Group III: 14.49, 24.68, 32.00, 32.97, 6.30, 50.72, 33.35
Group IV: 11.41, 18.14, 27.23, 45.97, 30.23, 30.58, 11.26

NON-TARGET CONSTRUCT 'TREATMENT' CONSISTENCY (C_1) SCORES

High Intensity Subjects:

Group I: 54.58, 50.15, 73.31, 79.91, 66.65, 40.27, 77.73, 54.88
Group II: 64.48, 53.23, 46.15, 24.96, 65.21, 54.31, 35.81, 36.81
Group III: 53.02, 55.32, 51.77, 47.53, 29.71, 57.96, 48.75, 38.09
Group IV: 51.50, 29.49, 30.87, 38.38, 6.22, 80.69, 28.15, 24.69

Low Intensity Subjects:

Group I: 38.63, 49.07, 36.24, 21.77, 26.40, 54.42, 27.43
Group II: 41.14, 7.25, 54.53, 48.08, 32.63, 59.09, 43.37
Group III: 24.27, 44.57, 49.53, 42.91, 21.58, 48.53, 19.53
Group IV: 43.08, 20.95, 45.93, 58.74, 53.93, 44.47, 31.02

NON-TARGET CONSTRUCT 'POST-TREATMENT' CONSISTENCY (C_1) SCORES

High Intensity Subjects:

Group I: 65.63, 63.32, 50.85, 68.75, 87.78, 26.56, 56.35, 43.09
Group II: 53.65, 33.40, 53.05, 41.73, 60.99, 46.65, 31.84, 27.57
Group III: 56.28, 39.03, 57.09, 48.86, 39.21, 64.84, 80.01, 41.71
Group IV: 32.55, 31.78, 43.86, 44.32, 25.72, 57.16, 34.43, 8.35

Low Intensity Subjects:

Group I: 39.81, 31.31, 24.94, 65.30, 48.22, 32.11, 13.72
Group II: 32.61, 12.37, 59.84, 57.87, 46.51, 48.37, 16.35
Group III: 18.79, 27.70, 41.70, 37.72, 26.32, 32.32, 22.88
Group IV: 52.65, 23.83, 42.07, 65.14, 27.27, 59.99, 17.89

TARGET CONSTRUCT 'TREATMENT' CONSISTENCY (C_{1-3}) SCORES

High Intensity Subjects:

Group I: 78.66, 74.75, 65.89, 82.67, 76.02, 62.89, 82.01, 60.29
Group II: 75.43, 70.63, 72.72, 47.64, 87.21, 78.70, 66.71, 62.52
Group III: 85.43, 64.72, 67.10, 49.85, 40.38, 80.36, 74.48, 62.86
Group IV: 40.28, 49.76, 66.21, 73.11, 67.35, 75.53, 55.75, 59.03

Low Intensity Subjects:

Group I: 60.04, 63.52, 74.13, 50.12, 71.74, 57.46, 84.49
Group II: 60.12, 61.28, 58.76, 61.14, 69.70, 58.05, 73.89
Group III: 61.08, 67.49, 70.78, 80.14, 50.30, 68.46, 56.28
Group IV: 54.05, 60.70, 70.48, 66.78, 76.31, 78.90, 51.61

TARGET CONSTRUCT 'POST-TREATMENT' CONSISTENCY (C_{1-3}) SCORES

High Intensity Subjects:

Group I: 82.73, 82.00, 82.27, 96.49, 64.66, 49.35, 64.56, 65.54
Group II: 62.34, 52.42, 78.95, 57.74, 75.55, 74.81, 78.12, 50.01
Group III: 78.03, 64.51, 62.06, 43.16, 46.08, 64.10, 70.06, 74.69
Group IV: 31.02, 57.00, 72.40, 67.88, 58.64, 63.11, 50.04, 38.04

Low Intensity Subjects:

Group I: 49.51, 67.16, 81.18, 66.57, 71.09, 47.68, 76.78
Group II: 67.67, 68.07, 57.52, 72.06, 88.37, 68.12, 77.34
Group III: 61.95, 81.34, 69.89, 53.31, 49.34, 82.95, 63.11
Group IV: 62.48, 52.28, 58.71, 76.84, 66.03, 54.13, 47.81

NON-TARGET CONSTRUCT 'TREATMENT' CONSISTENCY (C_{1-3}) SCORES

High Intensity Subjects:

Group I: 69.45, 79.30, 89.77, 90.17, 82.78, 61.31, 89.58, 81.95
Group II: 83.68, 82.50, 65.89, 53.61, 81.18, 73.67, 60.21, 79.70
Group III: 82.43, 84.63, 68.74, 68.54, 74.28, 85.25, 80.68, 82.59
Group IV: 69.14, 34.59, 73.14, 62.42, 41.08, 95.42, 49.92, 54.24

Low Intensity Subjects:

Group I: 78.73, 65.98, 74.71, 65.65, 67.99, 76.24, 55.75
Group II: 73.70, 64.43, 78.00, 72.09, 62.81, 74.26, 68.04
Group III: 53.57, 77.72, 68.03, 82.02, 70.23, 78.48, 70.17
Group IV: 70.02, 47.63, 69.79, 69.83, 79.89, 80.52, 56.23

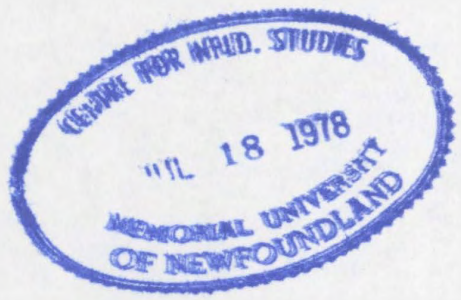
NON-TARGET CONSTRUCT 'POST-TREATMENT' CONSISTENCY (C_{1-3}) SCORES

High Intensity Subjects:

Group I: 83.20, 79.17, 67.75, 86.32, 94.80, 57.18, 82.20, 72.54
Group II: 83.05, 65.75, 66.74, 69.86, 80.54, 67.46, 77.57, 70.30
Group III: 73.64, 69.44, 71.91, 62.48, 58.95, 79.60, 80.01, 74.29
Group IV: 57.07, 52.65, 68.61, 62.98, 54.10, 82.33, 55.76, 40.49

Low Intensity Subjects:

Group I: 66.79, 53.16, 56.87, 74.63, 81.13, 67.50, 64.32
Group II: 74.68, 50.41, 78.72, 72.99, 84.34, 67.31, 44.29
Group III: 40.59, 73.01, 76.33, 55.84, 79.06, 73.43, 52.84
Group IV: 77.43, 48.08, 58.49, 71.29, 62.73, 76.90, 52.07


GRENFELL HIGH MARIT. STUDIES
MAR 18 1978
MEMORIAL UNIVERSITY
OF NEWFOUNDLAND