71-23,708

BRAVER, Sanford Lawrence, 1943-
REAPPRAISAL OF THE SELF THROUGH KNOWLEDGE OF THE PERFORMANCE OF OTHERS: IDENTIFICATION VS. SOCIAL COMPARISON PROCESSES.

The University of Michigan, Ph.D., 1970
Social Psychology

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

REAPPRAISAL OF THE SELF THROUGH KNOWLEDGE OF THE PERFORMANCE OF OTHERS: IDENTIFICATION VS. SOCIAL COMPARISON PROCESSES

by
Sanford Lawrence Braver

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
1971

Doctoral Committee:

Professor Melvin Manis, Chairman
Professor Eugene Burnstein
Professor John R. P. French, Jr.
Professor Alvin Zander

PLEASE NOTE:

Some pages have small
and indistinct type.
Filmed as received.

University Microfilms

ABSTRACT

REAPPRAISAL OF THE SELF THROUGH KNOWLEDGE OF THE PERFORMANCE OF OTHERS: IDENTIFICATION VS. SOCIAL COMPARISON PROCESSES

by
Sanford Lawrence Braver

Chairman: Melvin Manis

Part of the burden of this paper was to point out a conflict in the predictions of two theories, each rich in supporting literature. When a person identifies with another, according to the concept of identification, he vicariously shares the other's experience and reactions. In the case of self-appraisal, the other's successes will be accompanied by increases in the person's self-appraisal; the other's failure will evoke a decline in the person's self-appraisal. The notion of social comparison postulates that people compare their attributes to those of referent others in order to evaluate their own, and desire comparisons favorable to them. In the case of self-appraisal this implies that the more successful the other's performance, the more will the person suffer diminution in his self-appraisal; the worse the other's failures, the greater will be the person's self-enhancement. When the other is at once a referent and an identificatory other, the two

processes come inescapably into conflict. When the other succeeds, social comparison predicts the individual will lower, identification predicts he will raise his self-appraisal. When the other fails, the theories predict the opposite, still antagonistic, outcome.

Several variables, both situational and personality, were postulated to influence the resolution of this conflict, that is, to determine which of the counterpoised processes would prevail. The situational variables were: (1) The certainty with which the individual knew of the worth of his own performance; it was hypothesized that the greater the certainty, the more social comparison effects. (2) The nature of the competition between the person and the other; it was hypothesized that the more other was regarded as a rival rather than as an ally, the more social comparison effects. The personality variables were: (3) The strength of the identificatory bond. The hypothesis was that the greater this variable, the greater both effects. (4) The initial level of the individual's self-esteem; and (5) his field-independence. In the latter two cases it was predicted that the lower the variable, the greater the identification effects.

A 2 x 2 x 2 factorial experiment using grade-school subjects, and the bond of sex-role identification, was performed. After completing measures of self-esteem, strength of identification and Field-Dependence, and taking a bogus task, Ss were informed either that the other boys

in the class had performed well or poorly on the task. Also varied was the nature of competition between the subject and his male classmates, either Between Team (allies) or Within Team (rivals) Competition; and whether or not the subject was given feedback for his own task performance, either Score or No Score. The measure of identification and social comparison effects was the direction and magnitude of self-reappraisal, as measured by change in self-esteem, evaluation of the tested ability, and two allied measures.

The results showed that the instruments and manipulations were effective. However, randomization failed to equate the conditions on initial level of self-esteem, nearly nullifying the primary dependent variable, change in self-esteem. The measure of evaluation of the tested ability, however, indicated that the situational variables influenced the resolution of the conflict in the hypothesized fashion. Of the personality variables, only field-dependence reacted with self-appraisal substantially as hypothesized.

In addition there were several unpredicted significant effects. In each case, the post-hoc interpretation which was proposed was not totally congruent with additional data available to test it.

Suggestions for future research, in part stimulated by the serendipitous findings, were given. It was also speculated that the atypical nature of the sample was in part responsible for the several failures to replicate previous studies.

PREFACE

The development of the ideas for this dissertation has a long history. In 1964, two different papers I prepared as class assignments for undergraduate psychology courses actually formed the nucleus of the concepts employed here. One of these was the social comparison study reported herein; the other was an analysis of charismatic leadership, where I first began thinking about the notions of identification. The comments of Drs. Sandor Brent and Michael Schwartz were helpful at that time.

Sometime later, I recognized the relationship of the two studies to one another and began thinking along the lines of the present paper.

During the incubation and writing of these ideas, I benefited from discussions with many fellow graduate students, among them John Stevenson, Doug Cornell, Teri Levitin, Larry Schulman, Cy Ulberg and Malcolm Brenner. A student seminar on vicarious processes attended by John Stevenson, Malcolm Brenner, Pete Archibald, and Diane Kravitz also clarified my thinking.

Many members of the faculty not on my dissertation committee have contributed useful suggestions and I gratefully acknowledge the aid of Joe Veroff, Bob Zajonc,

Irwin Katz, Marty Gold, Keith Smith and Warren Norman.

To the students, parents, teachers and officials of the school system whose cooperation made subjects available, I am greatly indebted.

I appreciate, moreover, the aid of those counselors at the Statistics Laboratory and the Computer Center who gave patient advice when data problems were severe. Useful suggestions were provided in addition by Profs. Gary Galbraith and John Reich at Arizona State University.

I would like to extend heartfelt thanks to the members of my Committee, Drs. Burnstein, French, and Zander for their many intelligent and skillful suggestions. Often I was directed to more fruitful lines of approach by their useful comments. I must include in this acknowledgement the aid of Ted Newcomb who was a former Committee member.

Mel Manis is a doctoral candidate's dream as a chairman. His encouragement and wisdom were indispensable. He unstintingly provided as much incisive comment and moral support as could be asked. Some writers of dissertations seem to work better the more they are independent of their chairman. In contrast, I relied heavily on Mel, and he never failed to give timely advice.

To my wife, Judith, I give both thanks and love. She was a constant source of ideas, encouragement and assistance. During the work on this dissertation, my

identificatory bond with her, impossible though it seemed, strengthened. And my self-esteem increased at the renewed realization of the qualities of the girl I married.

Finally, I wish to thank my small son, Todd Samuel. His joyful cheer and abundant charm always coaxed my heart into smiles, thereby making enjoyable my labors.

TABLE OF CONTENTS

CHAPTER I. INTRODUCTION AND REVIEW OF THE LITERATURE 1

CHAPTER II. SYNTHESIS AND DEVELOPMENT OF HYPOTHESES 56

CHAPTER III. METHOD 79

CHAPTER IV. RESULTS - AND THEIR INTERPRETATION . . 102

CHAPTER V. CONCLUSIONS 160

REFERENCES 175

APPENDIX 186

LIST OF TABLES

Table		Page
1.	A Summary of the Hypothesized Effects of the Processes and Experimental Conditions on Reappraisal of Self	69
2.	A Summary of the 2 X 2 X 2 Factorial Conditions	70
3.	Classification of the 50 Items on the SEI (Coopersmith, 1968) According to the Direction of Keying and the Goodness of Item Reliability	87
4.	The Trait-Adjectives Used on the Affective Identification Instrument and Their Evaluative Ratings, as Given by Anderson (1964)	90
5.	KR-20 Coefficients of the Session II SEI for the Several Treatment Conditions	104
6.	Correlation Coefficients of Session I (Pre-) to Session II (Post-) SEI Scores, by Condition	105
7.	Analysis of Variance Summary Table for the Dependent Variable Team Success	111
8.	Analysis of Variance Summary Table for Initial SEI Scores	114
9.	Mean SEI, Version I, Version II, and Mean Gain (II-I), by Condition	116
10.	Analysis of Variance Summary Table for Post-SEI Scores	117
11.	Analysis of Variance Summary Table for Gain Scores as Computed by Special Formula Removing Regression Effects	122
12.	Analysis of Variance Summary Table (No-Score Conditions Only) for estimated score . . .	128

Table Page

13. Mean Estimated Errors, No-Score Condition Only, by Condition, and Marginals 128

14. Analysis of Covariance Summary Table for A-Eval, Utilizing Initial SEI as Covariate 131

15. Cell Means for A-Eval, Adjusted for Differences on the Covariate Initial SEI, by condition 131

16. Analysis of Variance Summary Table for S's Perceived Importance of Task 139

17. Correlation Coefficients for Strength of Identification with A-Eval (Hypothesized sign in Parenthesis) by Condition 151

LIST OF ILLUSTRATIONS

Figure		Page
1.	Means of SEI Change, Computed by Special Formula, Collapsing over Non-Significant Component, Team Success Failure	123
2.	Means for the Perceived Importance of Doing Well, After Collapsing on the Non-Significant Component, Score	140
3.	Means, Adjusted for Variation on Version I SEI, for the Perceived Importance of Doing Well, After Collapsing on the Non-Significant Component, Score	143

LIST OF APPENDICES

Appendix		Page
A.1	Letter to Parents of Subjects	186
A.2	Memo to Teachers of Subjects	187
A.3	Oral and Written Instructions and Instrument for Initial Measurement of Self-Esteem . . .	189
A.4	Oral and Written Instructions and Instrument for Measurement of Cognitive Identification	191
A.5	Oral and Written Instructions and Instrument for Measurement of Affective Identification	194
A.6	Oral and Written Instructions and Instrument for Measurement of Field-Dependence	196
A.7	Oral and Written Instructions and Instrument for Completion of Bogus Task	200
A.8	Instrument for Final Measurement of Self-Esteem	207
A.9	Supplemental Questions Asked of Subjects in All Experimental Conditions	209
B.1	Rationale and Derivation of Special Formula Used to Estimate a Subject's True Change	211

CHAPTER 1

INTRODUCTION AND REVIEW OF THE LITERATURE

Introduction: Reappraisal of the self through knowledge of own performance.

If man has engaged in a quest to "know thyself," so have psychologists attempted to determine how this self-knowledge is won. The question of the antecedents of self concept has engaged a substantial portion of the research efforts in the behavioral sciences. By far the predominant orientation informing this inquiry is the Mead (1934)-Cooley (1922) symbolic interaction hypothesis, which asserts that one's perception of the attitudes that significant others (e.g., parents, close friends, teachers) hold toward self shapes one's self-appraisal (Miyamoto and Dornbusch, 1956; Manis, 1955; Couch and Murray, 1964; Helper, 1958; Rosenberg, 1963; Videbeck, 1960).

A second approach which has received considerable attention concerns one's everyday success and failure experiences, by which information is achieved about one's capabilities and worthiness. The operational arrangements which afford a test of this notion generally employ an experimental task and feedback by an expert or objective

judge regarding the quality of S's performance on the task. Haas and Maehr (1956), for example, demonstrated not only that upward changes in self-esteem follow success and downward changes follow failure, but also that the changes are relatively enduring. A study by Sharma (1956) also found this effect. Dittes (1959), rather than using self-esteem as his dependent variable, employed a supposed concomitant of these changes, namely S's impulsive attempts to obtain closure on subsequent ambiguous tasks. Presumably such attempts, if successful, can restore a damaged self-esteem, and hence, there will be more perseverance if S's self-esteem has received previous injury than if it were intact. Granting these assumptions, his general finding that such perseverance followed previous failure more readily than previous success can be taken as further evidence that changes in self-esteem follow successful and failing performances.

Several issues arise with regard to this literature. The question of whether the theoretical mechanism at work here is distinct from that proposed by symbolic interactionists (since the feedback of success and failure is normally given by a significant other, an expert) and hence can be subsumed by the latter, need not overly concern us here. By contrast, the issue of the specificity of the change in self-appraisal--the spread of effect through the self-system--has implications for the present analysis.

We shall refer to global self-esteem (or simply self-esteem) when we intend the person's evaluation of himself as a totality. Self-evaluation (or ability-evaluation) shall be used to mean the person's evaluation of his standing on the specific trait or ability being tested. The term self-appraisal shall be used in the generic sense, as either type of evaluation. At issue, here, is whether a single success-failure experience with regard to a single, specific ability generalizes or radiates to produce changes in self-esteem, or simply produces changes in self-evaluation. The studies of Haas and Maehr (1965), Maehr, et al. (1962), Diggory and Magaziner (1959), Diggory (1966), Rothman (1963), and Cetlin (1964) are relevant here, and all give evidence for radiation or spread of effect. To be sure, the specific ability tested will demonstrate the largest shifts; related, non-tested abilities will show smaller changes; and the change in evaluation of unrelated abilities and global self-esteem will be smaller still, as Wylie in her review (1961) notes. The very conception and measurement of self-esteem make this finding tautologically true, since, in general, the index consists of summing the evaluations of specific, important abilities. Generally speaking, then, any event which affects the self-evaluation of an important ability will affect, to a lesser extent, global self-esteem.

A second issue which is of relevance for the present report concerns the permanence of change. Does this

single, specific success or failure experience have enduring effects on self-appraisal, or is it entirely transitory in nature? To the degree that it is the latter, does it then become trivial? Perhaps the best answer to the second question is that most if not all instances of self-appraisal are somewhat transitory. The weight of all previous learning about the capacities of self, plus all the learning which comes after--and this learning is virtually continuous--easily overwhelms a single bit of information. The effects, if fleeting, are so only because later learning, and assimilation of this information to previous learning, reduce its salience over time. If all effects on self-esteem are more or less transitory, and we don't regard self-esteem as trivial, then we cannot so regard transitory influences.

The key fact here is the degree to which it is transitory. If the ability being tested were regarded as very important, and there had been but infrequent opportunity to evaluate it in the past or future, then the effects of success or failure on self-appraisal would, without doubt, be relatively enduring. In the Haas and Maehr study (1965) for example, the changes lasted over a considerable period of time, and faded more quickly for self-esteem than for self-evaluation. In a study by Hicks (1962), however, both self-esteem and self-evaluation remain elevated by success as much as five days later.

Thus, at least some success-failure experiences are found to have long run effects.

The Performance of Others

Granted, then, that the evidence indicates that self-appraisal is affected by knowledge of the aptitudes we display in the performance of tasks, we can now proceed to the inquiry which impels this investigation: can self-appraisal be affected by knowledge of the performance of others?

To start the discussion, let us reflect on an example which demonstrates the process. A novice tries his hand at boxing, and in his first bout, scores a victory. His appraisal of his boxing ability is still somewhat ambiguous, but consider the effect of learning his opponent had been a Golden Gloves Champion. In this instance, the knowledge of his opponent's past performance would have clear implication for S's self-appraisal. (Note that whether the novice won or lost the fight, knowledge that his opponent was good enhances self-appraisal, knowledge that his opponent was inferior deflates it.)

This particular mechanism, while it serves amply to convince the reader that knowledge of the performance of another can promote reappraisal of the self, is not notably theoretically rich. There are, however, two social psychological processes which have been extensively researched, which similarly suggest themselves as mediators

between knowledge of the performance of others and re-appraisal of the self. The two are social comparison processes and identification processes. In the next section, we will review the literature regarding the process of identification.

Identification

Consider the following examples: (1) a young boy, learning that his father won the club's golfing tournament; (2) a father, learning that his son got accepted to Harvard; (3) a wife, learning that her husband can't hold a job; (4) a New Yorker, learning that the Mets won the World Series; (5) an American, watching a countryman walk on the moon; (6) a Black youngster, learning for the first time that some people of his race were notable and respected in American history. In addition to experiencing a feeling of joy, the subjects in these examples probably undergo changes in their appraisals of themselves. They become proud--proud of <u>another</u> rather than themselves, to be sure--but proud nonetheless. And to feel pride <u>means</u> to experience a sense of worthiness. The wife in example 3 probably feels shame, a diminuation of self-esteem, as a consequence of the knowledge of her husband's failure. As a general statement, it can be said, that <u>in some circumstances</u>, people will share the self-appraisal of the performing other: they will gain in self-esteem when the other succeeds, just as he will; they will decrease their self-appraisal when he fails, also just as he will.

The most apparent circumstance in which the above result obtains, of course, revolves around the nature of the relationship between the subject and the person whose accomplishments he learned of. In each of the examples, this relationship was of a special nature: father-son, husband-wife, sports fan, etc. The sharing of the self-appraisal, in all probability, would not occur if this relationship did not exist--if the subjects learned of some stranger's performance, for example. We shall be careful to call this kind of relationship one in which identificatory bonds exist, or simply identification; we shall distinguish this from the occurrences of sharing of self-appraisal which we shall call the identification effect, process or hypothesis. We shall find that this usage is to a degree distinct from the way the concepts have been used traditionally in the past literature.

Theorists of Identification: Freud and Mowrer

In reviewing this literature, the reader should be warned against expecting that a unified whole will emerge. Different writers, each from their own vantage points, speak to different aspects of the phenomenon. Empirical research has little overlap with--does not extend from--theory. The result approaches a hodge-podge of literature, upon which attempts to impose organization will generally be superficial.

The term identification was first used by Freud in

1917 who later used the concept to explain the fundamental problem of socialization, the development of a conscience or superego, of internal controls over behavior. Noting that there is a tendency for children to imitate their elders, he formulated the notion that this imitation was but an observable manifestation of a generalized tendency to become like their parents--to swallow them and become them, to take on their identity. As Bronfenbrenner (1960) puts it:

> In concerning himself with identification Freud was not asking why and how a child might learn an isolated piece of behavior from his parent. He was interested in what he felt to be a more sweeping and powerful phenomenon--the tendency of the child to take on not merely discrete elements of the parental model but a total pattern. Moreover, as Freud saw it, this acquisition was accomplished with an emotional intensity which reflected the operation of motivational forces of considerable power. These features of pattern and power are reflected in his use, as virtual synonyms for identification, of such terms as introjection and incorporation--words which connote a sudden and somewhat desperate "swallowing whole" of the parent figure. (pg. 27)

Thus, the formerly external forces (i.e., the parents) containing the child's instincts become internalized via this process.

The mechanism, the interpersonal sequence which forces this identification, was quite complex but of integral importance for the theoretical fabric of psychoanalysis. For the young male, the process is postulated to be as follows: having developed with the mother an object-relationship, due in part to his experience at the

breast, when he enters the phallic stage at the age of two or three, he desires to become her lover.

> He desires to possess her physically in the ways in which he has divined from his observations and intuitive surmises of sexual life and tries to seduce her . . . In a word, his early awakened masculinity makes him seek to assume, in relation to her, that place belonging to his father, who . . . now becomes a rival who stands in his way and whom he would like to push aside. (Freud, 1949b, pg. 91)

However, he begins to fear that his father will castrate him in retaliation for his rivalry. To resolve this Oedipal dilemma, he renounces his sexual love for his mother and identifies with his father:

> Frightened by the overwhelming threat of castration, the little boy accepts the fact that he will be vanquished by the father. He realizes that he cannot compete successfully for the mother's love. Identification with the father then develops, replacing the boy's jealous, hostile, and rivalrous feelings toward him. It is as though the boy says, "If I am him, he cannot hurt me." Identification with the father then serves two functions for the boy: (1) it reduces the boy's fear of the father, and (2) it permits him to enjoy the mother's love vicariously. (If he <u>is</u> the father, he possesses the mother's love.) (Mussen, 1967, pg. 82)

As involved as this account is, it becomes even more complex in explicating the development of the little girl's identification with the mother. Since she has no penis, and hence no fear of castration, Freud found he needed a separate mechanism to account for how she comes to identify. This mechanism Freud termed anaclitic, and was motivated by fear of the loss of the mother, or the mother's love.

Fear of castration is naturally not the only

> motive for repression; to start with, it has no place in the psychology of women; they have, of course, a castration complex, but they cannot have any fear of castration. In its place, for the other sex, is found fear of the loss of love, obviously a continuation of the fear of the infant at the breast when it misses his mother. You will understand what objective danger-situation is indicated by this kind of anxiety. If the mother is absent or has withdrawn her love from the child, it can no longer be certain that its needs will be satisfied and may be exposed to the most painful feelings of tension. (Freud, 1933, pg. 121)

In his Outline of Psychoanalysis, Freud elaborated on the consequences of this fear of the mother's withdrawal:

> . . . the daughter, under the influence of her envy for the penis, cannot forgive her mother for having sent her into the world so insufficiently equipped. In her resentment, she gives her mother up and puts someone else in place of her as the object of her love--her father. If one has lost a love-object, the most obvious reaction is to identify oneself with it, to replace it, as it were, from within, by means of identification: this mechanism now comes to the little girl's assistance. Identification with her mother can take the place of attachment to her mother. The little girl puts herself in her mother's place, as she has always done in her games; she tries to take her place with the father . . . (Freud, 1949a, pg. 98)

One critic (Stoke, 1954) has been particularly stinging in his attack on the Oedipal explanation of the development of identification:

> Its etiology is particularly vulnerable for the Oedipus complex is not scientifically verifiable, is inconsistent in its application to boys and girls, fails to account for differences as well as similarities between generations . . . , emphasizes organic causes of identification to the virtual exclusion of others, and is not very useful in child guidance. (pg. 229)

It is clear that, however convoluted Freud's explanation of the process, the consequences of the identifica-

tion is an emotional cathexis or tie to the parent, an alignment of self with other. While other writers disagree with Freud about the evolution of identification, most agree that the ultimate bond, or sense of attachment that develops is a genuine product of child development which in turn is a concept necessary to explain child behavior. Bandura (1969), however, is one who does not agree. In terms of observable behavior, he says, identification is simply "the tendency for a person to reproduce the actions, attitudes, or emotional responses exhibited by real life or symbolized models." And this is precisely how he defines imitation. Thus, he objects to distinguishing the two:

> Essentially the same learning process is involved regardless of the context in which it is learned, the object from whom it is learned, or the stimulus situations in which the relevant behavior is emitted. Therefore, it is in the interest of clarity, precision, and parsimony to employ the single term, imitation, to refer to the occurrence of matching responses. (Bandura and Walters, 1963, pg. 89-90)

One could argue with Bandura that he has missed the point. Where "the occurrence of matching responses" can be seen to vary solely as a function of the relationship between the identifier (henceforth termed S, the subject) and the identificand (henceforth the Model, M)--and we shall later present evidence even from Bandura's own work that it does--appeal must be made to an intervening variable to explain the findings. We have labeled this variable identification. While other writers agree that the

nature of the relationship of S to M is crucial for the distinction between identification and imitation, some of these have persisted in defining the former as response matching with the proviso of attachment of S to M, instead of as the attachment per se, as is done in the present paper:

> The two major mechanisms of the learning of patterns from social objects are imitation, which assumes only that alter provides a model for the specific pattern learned, without being an object of generalized cathectic attachment; and identification, which implies that the alter is the object of such an attachment and therefore serves as a model not only with respect to a specific context of learning, but also as a model in a generalized sense. (Parsons and Shils, 1961, pg. 129)

This reluctance to relinquish imitative behavior as one of the defining characteristics of identification is no doubt an understandable attempt to provide an observable underpinning to a hypothetical construct. However, when it is noted that in bonafide instances of identification other than parental identification, such as the parents identifying with ("living vicariously through") the child, overt imitation of M's responses rarely occurs, this reluctance becomes self-defeating.

Mowrer, in a seminal paper (1950), recognized the distinctions in Freud's usage of identification for boys and girls, acknowledged the epistemological status of both, and proceeded to elaborate the concepts, which he now called defensive and developmental identification, respectively. Developmental identification, he says, is the

more normal pattern, can occur for boys or for girls, and is instigated by the child's desire "to reproduce bits of the beloved and longed-for parent" (pg. 615). By becoming the parent, the child maintains his presence, even when he is physically absent. It is marked by the dependence of the child on the parent, for then the motivation to replace the withdrawn object suffices to promote identification. A learning theorist, Mowrer argued that insofar as the nurturance and love of the parent is reinforcing, self-reproduction of the parent's behaviors takes on a secondary reinforcing quality.

Defensive identification, according to Mowrer, results when the parents are harsh, authoritarian, and threatening, and is not limited to the boy's fear of castration at the hands of his father. If the child becomes his threatener, he need no longer fear him. In the words of Jones and Gerard (1967), there is "a magical fantasied self-protection" brought on by the child's desire to reduce the threat inherent in his relative powerlessness.

In further distinguishing the two forms of identification, Mowrer writes:

> It is true in both developmental and defensive identification that the subject is "frustrated," but the different nature of the frustration in the two cases is noteworthy. In one case, it arises from a sense of helplessness and loneliness: the parent or parent person is absent and the infant wishes he were present. In the other case, the frustration arises rather from interference or punishment; the parent or parent person is

> present and the infant wishes he were absent. But the latter wish brings the average child into intolerable conflict: while he hates the parent for disciplinary actions, he also loves the parent and experiences acute anxiety at the prospect of his really being separated, physically or emotionally from him (or her).
>
> Developmental identification, we may suppose, is a milder and simpler experience than is defensive identification, which has a violent crisis-like nature. The one is powered mainly by biologically given drives ("fear of loss of love," in the analytic sense) and the other by socially inflicted discomforts ("Castration fear," or less dramatically, simply fear of punishment). The first presumably involves relatively little conflict; but in the latter case, conflict and dependent anxiety are outstanding. (pg. 572)

Since modern theorists are prone to believe that identification occurs primarily with an M that S admires, or feels positively toward (who gives rewards), yet Mowrer and Freud argue for the existence of identification with a disliked or a feared object (who gives punishment) as well, it would be well to examine the research which attempts to demonstrate the latter.

Sarnoff (1951) used Bettelheim's (1943) accounts of Jewish concentration camp prisoners taking over the characteristics of their Nazi tormentors as an impetus for his study. He reasoned that if a person has harsh, authoritarian parents, he will adopt a defensive form of identification with them. This mode will then characterize all later identifications with others. In particular, the Jewish child with authoritarian parents should prefer defensive identification, that is identification with those who

threaten him, namely anti-Semitic gentiles. He should come, in his identification, to take on the characteristics of the prejudiced Christians, including their anti-Semitic views. This hypothesis was confirmed. Similarly, anti-Semitic Jews should have lower self-esteem, since they agree with their tormentors about themselves as well; and should habitually respond passively toward aggression directed toward them. Both of these hypotheses were confirmed also.

Maliver (1965), however, tested these same hypotheses using Black students' anti-Negro feelings of self-hatred, and found that all three hypotheses failed to be confirmed.

Baxter, Lerner and Miller (1965) divided Ss, on the basis of pre-experimental questionnaires, into those whose home socialization experiences were severe vs. democratic. These Ss then participated in a learning task where the E either gave reinforcement for correct responses, punishment for incorrect responses, or, for the control sample, merely knowledge of results. After this experience, Ss provided self-descriptions and descriptions of E, as well as that of "the average person." After partialing out the latter from the correlation of the first two, they found that Ss of severe home-life attributed self-like characteristics (their measure of identification being assumed similarity) most to the shocker, while Ss of nurturant home-life described the rein-

forcing E as most like themselves. Providing, then that one doesn't quarrel with their measure of identification, they demonstrate both identification with the aggressor, and developmental identification, depending on S's prior relationship with his parents.

Carlson (1963) attempted to pit two hypotheses against one another: (1) as strength of identification with parents increases, so does self-esteem; (2) self-esteem is higher for those whose parental identification is based on a supportive relationship than for those for whom it is based on threat and power. To the degree that the latter and not the former hypothesis is true, she argued, support for the distinction between types of identification--and the existence of defensive identification--is established. She asked her pre-adolescent Ss and both parents to describe the ideal child, and the way S actually was. Assuming that the degree of internalization of parental standards was a consequent of strength of identification, she defined the latter as congruence between the parents' and child's description of the ideal child. Type of identification was defined as defensive if any of the following four conditions were met, developmental otherwise: (1) if parents' descriptions of their child were severely different from that of their ideal child; (2) if parents' descriptions of the child were severely different from the child's self-description; (3) if the child's description of the ideal child was closest to that of the

least supporting parent (i.e., the parent that described his child as deviating most from his ideal child); (4) if the child's ideal differed severely from that of the parents. Utilizing these measures, she affirmed that type of identification was of greater consequence than strength of identification in influencing self-esteem.

At face value, all but the Maliver (1965) study support the notion of identification with the aggressor. However, all may be faulted on methodological grounds. To pick an example almost at random, in the Baxter et al. study (1965), the measure of identification is highly suspect. If we assume that persons raised in a severe environment have lower self-esteem, they would describe themselves unfavorably, relative to those from a secure environment. Moreover, persons subjected to a rewarding E should like him more, describe him more favorably, than Ss exposed to the punishing E. Thus nurturantly raised, rewarded Ss like themselves and E; while authoritarianly raised, punished Ss dislike both themselves and E; in both cases giving similar evaluations, assuming similarity between self and E. These two plausible assumptions account for the data without invoking the identification concept at all. Thus the evidence in favor of the existence of the phenomenon of identification with the aggressor, it must be concluded, is spotty, at best.

In the hands of later writers, as Sanford (1955) points out, the term identification is defined in many

different ways, nearly all of them vague; it is not clearly and consistently distinguished from empathy, vicarious living, sympathy, altruism, closeness, loyalty; it can refer to identification with other people, singly, or in groups, with causes, with animals, etc.; it can be expressed in fantasy, conscious experience or overt behavior. Indeed, he continues, "a term that can be employed so many different ways, and that, as Tolman says, has been accepted by most psychologists and sociologists, could hardly mean anything very precise. It might be proposed, quite seriously, that we give up the term identification altogether." (pg. 107).

Most theorists of identification identify themselves as students of child development, (e.g., Sears, Stoke, Bronfenbrenner, Lynd, Kagan) and hence use as a model to formulate hypotheses the parental identification of children. The difficulty with this strategy is that whereas each kind of identification we have referred to has elements in common with the others (else all could not be exemplars of the same concept) they also each have a number of unique characteristics. For example, the identification of a wife with her husband differs in myriad ways from the identification of a Black man with his race. To try to draw conclusions about all these forms, using a perspective as narrow as any single one, such as parent-child identification, is bound to be hazardous and misleading.

It is probably possible to place recent theorists roughly along a continuum according to the degree to which learning principles are central to their conceptions. Stotland, an almost purely cognitive theorist, and Sanford (1955), one of psychodynamic orientation, would both be placed at the low end; Bandura would define the highest end of the continuum; while Parsons (1955), Lynd (1969), Kagan (1958), Whiting (1960), Maccoby (1959), Stoke (1954), Sears (1967) and Lazowick (1955) would be intermediate, but not necessarily in that order.

Sanford (1955) used as his model a psychiatric patient's identification or transference with the therapist, arguing that the reverse case, the therapist's vicarious living through his patient, does not qualify for the term identification and should be subsumed by other processes. Since his perspective differs from that of the developmental psychologist, his characterization of the identification process is somewhat different. Whereas Mowrer and most other theorists see identification as a desirable, and, in any case, inevitable development, Sanford sees it as a maladaptive crisis reaction. He observes, "a measure of desperation in one who identifies," (pg. 109), as well he might, for there is a measure of desperation in anyone who seeks the services of a therapist. He concludes that the primary dynamic source of identification is a threat to self-esteem.

While Sanford is concerned primarily with the cir-

cumstances under which a person identifies, Whiting (1966), Kagan (1958) and Maccoby (1959) all formulate hypotheses about with whom S will identify. In all these hypotheses, M is superior in some way to S. To Whiting, M has superior status; Maccoby argues that M has resources that S needs; while Kagan emphasizes that M is obtaining satisfaction of his own needs, which accomplishment S envies.

By status, Whiting intends the degree to which M controls resources. If M readily gives these resources to S, no identification will develop, for then S will have control of them.

> He will, on the contrary, envy the status of those resource mediators who withhold resources from him, deprive him of resources that were formerly his, and consume or enjoy those resources in his presence. This process may be termed envy of resource mediator. If the resource mediator withholds a resource from a child and gives it to a third person, this third person will occupy the envied status. (pg. 118)

In contrast, Maccoby contended that it is not the person who consumes the resources, but the person who can decide how to dispense them that is crucial.

> If most things that an individual wants are not under his control but under the control of others, then presumably much of the vicarious trial and error that he engages in must involve his trying out various approaches to getting the help of, and avoiding the censure of others, and imagining the probable responses to these approaches . . . a child should covertly rehearse both the rewarding and the punishing actions characteristic of his parents. (pg. 245-246)

An experiment by Bandura, Ross and Ross (1963b)

tests these differing nuances. Nursery school children participated with two adults in game-like interaction. One of the adults always had control of the resources. For half the Ss, the other adult was the recipient or consumer of the resources, while the child looked on; for the other half, the roles of the second adult and S were reversed. The data consisted of the number of the adults' ideosyncratic behaviors that the child imitated when alone. The controlling adult was always most readily imitated, even when the second adult was allowed to consume the resources.

Mussen and Distler (1959) found that boys who achieve strong masculinity preferences, presumably those whose identification with their fathers was strong, characterized their fathers as more powerful--both more rewarding and more punishing--than boys with low masculinity preferences, a result that corroborates those of the Bandura et al. study. Mussen (1961) and Payne and Mussen (1956) found only the rewarding, not the punishing aspect of the above results.

Kagan (1958), Sears (1957), and Lazowick (1955), all used reinforcement principles extensively in their formulation but in distinct ways. Sears actually takes much of the blood out of the identification concept, by reanalyzing it solely in terms of secondary reinforcement. For him identification occurs in three developmental steps: (1) first the child develops a dependency drive "for which the mother's affectionate nurture is the approriate

environmental event;" (2) the child begins to imitate in rudimentary form the mother's behavior; (3) the imitation is found to be self-gratifying, so that the response tendency to imitate increases in strength. Thus, "identification is a secondary motivational system. The goal response is acting as if one possessed the psychological properties of another person." (pg. 156). Sears wishes the identification concept to explain only why one acts like another person when the particular imitative responses have not been previously reinforced.

Lazowick (1955) similarly restricts the meaning of identification to the learning of other's responses, but finds secondary reinforcement an insufficient explanation to cover the range and depth of imitative responses, such as similarity of attitude, etc. Hence, he invokes the concept of mediation. A student of Osgood's, he asserts that, in the course of socialization, the semantic space of the child comes to be shaped into the pattern of the parents; and that this similar way of viewing the world tends to result in the emission of similar responses. In testing his hypothesis, however, he found the semantic space of the child not to resemble that of either of his parents any more than it did that of an adult selected at random. Helper (1955), also a student of Osgood's, found essentially the same results, though he used the similarity of the child's ideal self to the parent's ideal child as his measure, rather than semantic space.

Kagan (1958) formulated one of the more interesting conceptions of identification. He defines identification as an "acquired cognitive response," in which "some of the attributes, motives, characteristics, and affective states of the M are part of the S's psychological organization. The major implication of this definition is that the S may react to the events occurring to M as if they occurred to him." (pg. 296). In this explication, we see that empathy, the vicarious sharing of experience, is a part of the definition of identification, a rather interesting development. Kagan posits an identification response ("some of M's characteristics are mine") which is subject to the usual principles governing the strength of a response. The scenario is as follows: initially S perceives that M possesses or commands goals and satisfactions that S desires. S assumes that the more similarities there are between S and M, the more likely S is to possess or command the desired goal state of M. Each time S perceives a similarity to M, the identification response is strengthened. Next, Kagan dealt with vicarious reinforcement. If S believes that M experiences positive affect (i.e., M's desired goal states), S experiences positive affect, which cycles back to further reinforce the identification response. On the other hand, if S perceives that M is suffering negative affect, the identification response extinguishes. The strength of identification is a joint function, says Kagan, of the motive to identify, and the

quality and quantity of reinforcements. S's perception that M has reached his goal states, it can be clearly seen, has two effects: it increases S's ability to experience M's reinforcement vicariously; it increases S's tendency to see himself as similar. Though Kagan's analysis is not terribly lucid, it is seminal in the sense that he uses the concept in a broad sense (the principles are not applicable only to that of parental identification) and he makes the concept of vicarious experience of another's psychological state a postulate of his system, from which he derives attempts to be similar, i.e., imitation.

Stoke (1954) similarly differentiated the bond through which vicarious affective states pass from imitation, though he did so in a less incisive fashion than Kagan. He termed the former emotional identification, and the latter behavioral identification, and proceeded to hypothesize about the relation between them. Emotional identification, he contends, produces a desire for behavioral identification, whereas lack of emotional identification will result in the absence of conscious attempts at behavioral identification.

Theorists of Identification with an Experimental Program: Bandura and Stotland

Of all the theorists, the positions of Stotland and Bandura are probably the most divergent from one another, yet each can claim, relative to the remaining theorists,

massive empirical support. Bandura argues for the simplification of the identification concept. What is really crucial, he observes, is whether S performs any of M's overt responses. Since a learning process explains these behaviors, it is rarely necessary to invoke such mystical concepts as identification. According to his theory, however, it is necessary to distinguish acquisition from performance of responses. As his 1965 study demonstrates, if a response is novel enough, it will be acquired. Whether or not S performs this response depends upon the reinforcement contingencies to himself and to M. Acquisition depends upon coding M's response in symbolic or ideational form. In a study he considers one of the more sensitive tests of his theory (Bandura, Grusec and Menlove, 1966), Ss were either (1) instructed to verbalize the actions of a filmed model as they occurred; (2) given no instruction; or (3) were instructed to give counter verbalizations (i.e., rapid counting, which impeded coding). Group 3 was least able to reproduce M's behavior, while Group 1 was most able, when given incentives to do so.

It is noteworthy that Bandura, while inveighing to abolish conceptual distinctions between imitation and identification, uses almost exclusively kindergarten (or younger) Ss in his empirical investigation of this concept. (Bandura, Grusec and Menlove, 1967a; Bandura and Harris, 1966; Bandura and Huston, 1961; Bandura and Kupers, 1964; Bandura and McDonald, 1963; Bandura and Whalen, 1966;

Bandura, Ross and Ross, 1963a and 1963b; Bandura and Mischel, 1965). The exceptions (Bandura and Rosenthal, 1966; Bandura, Grusec and Menlove 1967b; Bandura and Menlove, 1968) are generally studies of vicarious classical conditioning and extinction. This subject selection calls the generality of his position into serious question. If adults were used, of course, imitation would rarely be found. Even if M were to behave in an extraordinarily nurturant fashion, it is difficult to believe that an adult would imitate his novel responses to more than minimal extent. Yet as Bandura uses imitation and identification synonymously--or the former as the measure of the latter--this would imply that adults, who do not imitate as do children, cannot identify as do children. In children, imitation can be readily equated with the intuitive meaning of identification. In adults, however, their identifications are not often manifested in imitative behaviors, and their imitative behaviors rarely imply identification. They can identify with their spouse, but not imitate. When they do imitate--as in the instances of behavioral contagion--it is not because or with whom they identify. Before we, in tribute to parsimony, dispose of duplicatory concepts, prudence requires that one be shown unconditionally superfluous. In order for Bandura to do this, he will have, at least, to demonstrate that his arguments apply to adults as well as children. To the writer,

however, it seems that adult identification requires a different set of principles than those Bandura employed.

In order to discuss remaining contributions by Bandura, and, as well, the work of Stotland and his associates which we discuss next, it is useful to develop the concept of empathy and distinguish it from identification and imitation. As we shall use it here, empathy will mean the emotional or affective experience that a person undergoes as a result of the observation of the lot of another, (Heider, 1958) and the reactions of that other to his lot. We shall present data below that indicates that S's autonomic responses change as a function of M's circumstances, so that empathy as we define it here is not a null event. The importance of the concept, we shall say here however, is that the sharing of emotion--which for our purposes we need classify only crudely as having two dimensions: (1) intensity and (2) pleasantness or unpleasantness-- is a truly vicarious process. When a person feels joy because another feels joy, or dismay because another feels dismay, hedonic states are being shared. It then becomes almost a corollary to argue that self-appraisal my be shared as well, since the latter may be merely a special case of hedonic state, in that it involves "feeling good" or "feeling bad" about oneself or about one's abilities and characteristics.

It is interesting to note that Bandura's explanation of the mechanism by which "the behavior of observers can be substantially modified as a function of witnessing

other people's behavior and its consequences for them" (1969, pg. 30) relies upon the behavioral equivalent of the empathy concept. It illustrates that even a behaviorist must acknowledge that affective experience "flows" between people--that emotive experience is not confined "within the skin" but is also a "between skins" event. He first acknowledges several more mundane interpretations, for example, the fact that "response consequences experienced by another person undoubtedly convey information to the observer about the probable reinforcement contingencies associated with an analogous performance in similar circumstances." Then, however, he proposes a truly vicarious process, albeit described in S-R terms, as a necessary explanatory principle.

> A vicarious reinforcement event not only conveys information concerning probable reinforcement contingencies, knowledge about the types of situations in which the behavior is appropriate, and displays of incentives possessing activating properties, but it also includes affective expressions of models undergoing rewarding and punishing experiences . . . the pleasure and pain cues emitted by a model generally elicit corresponding affective responses in the viewer. These vicariously aroused emotional responses can readily become conditioned either to the modeled responses themselves, or to the environmental stimuli that are regularly correlated with the performer's affective reaction. As a consequence, the subsequent imitation of matching responses by the observer or the presence of the correlated environmental stimuli is likely to generate some degree of emotional arousal . . . It is therefore possible that the facilitative or suppressive effects of observing the affective consequences for the model may be partly

> mediated by the vicarious conditioning or extinction of emotional responses. (1969, pg. 31-32, emphasis mine).

Bandura bases much of his analysis on the classic Berger (1962) study in which Ss who were told that M was being shocked acquired conditioned fear responses (GSR) to the M's CS. Bandura's study with Rosenthal (1956) replicated and extended this finding. Bandura argues that many pathological phobic behaviors are acquired in this vicarious way: "Many phobic behavior arise not from actual injurious experiences with the phobic objects, but rather from witnessing others respond fearfully toward, or be hurt by, certain things." (1969, pg. 167). Thus, he reasoned, these same conditioned fear responses can also be extinguished vicariously. A series of studies puts this notion to empirical test. One reported by Bandura, Grusec, and Menlove (1967b) is illustrative. This well-controlled study took children with dog phobias, and in a series of eight sessions, presented a M interacting with the dog in increasingly fearful ways. At the conclusion of the sessions, each child was tested for dog phobia with either the same dog as the M played with, or a different dog. The substantial fear extinction that resulted was not specific to the M's dog, but generalized to any dog used.

Stotland's work on identification began when he attempted to use vicarious process to reduce prejudice (Stotland, Katz and Patchen, 1959; Stotland and Patchen,

1961). Ss read case histories of M which were designed to give self-insight into the psychodynamics of prejudice. The original study showed the expected reduction, but the follow-up found no reductions, but in some conditions actual increases in prejudice. In an attempt to formalize the notions inherent in these early studies, Stotland, Zander and Natsoulas (1961) posited the following simple principle: if S perceives a similarity to M on one attribute (First Similar Attribute or FSA), then S will tend to perceive similarities on other attributes as well (Derived Similar Attribute or DSA), provided that no contrary evidence is found. This tendency to generalize similarities was assumed to be one form of the phenomenon of identification. In the above study, Ss were told that they had an FSA in common with M; the results showed that they then predicted that M had a DSA that S did (projection); and that they, themselves, had a DSA that M did (introjection). In later studies, DSA was often evaluation of self on an ability. In the Burnstein, Stotland, Zander (1961) study, for example, M was either made to appear similar or dissimilar to S, and was made to appear expert or inept at deep sea diving. Though their data are to some degree equivocal, they nevertheless demonstrate that only when S is similar to M does S's evaluation of his own deep sea diving ability follow that of M. In the Zander, Stotland, and Wolfe (1960) study, rather than a single M, a group was formed, and identification with this

group manipulated by giving instructions which increased or decreased the group's unity. Self-esteem was significantly changed only when group identification (unity) was high, and when the group failed.

Other studies extended the original FSA⟶DSA formulation somewhat. Stotland and Dunn (1962) and Stotland and Hillmer (1962) introduced the concept of "Oppositeness": if S perceived he was opposite to M on FSA, then they hypothesized that he would tend to perceive oppositeness on DSA as well. The data supported this hypothesis only marginally, i.e., in a limited set of conditions. Also, these two studies demonstrated that the FSA⟶DSA process operates even when the FSA is arbitrary, even when it does not logically imply further similarity. For instance, in the Stotland and Dunn study, FSA was manipulated by informing S that M had been randomly selected to work on the same previous experimental task as did S. Even in this condition, the self-evaluation of S (as well as his subsequent performance of a later task) was influenced by how well he was told M performed. In a series of studies with Cottrell (1961, 1962; Stotland, Cottrell and Laing, 1960), the hypothesis that increased interaction with M can, as well as information regarding FSA, affect the imputation of DSA's, with regard to both introjection of self-appraisal, and projection, was tested, with mixed results.

In his more recent work (1969; Stotland and Dunn,

1963; Stotland and Walsh, 1963), Stotland has investigated the related concept of empathy, rather than self-appraisal, as an outcome of the FSA⟶DSA process. In fact, Stotland defines empathy as a subprocess of identification: "Empathy is a form of identification where the identifier acquires M's level of anxiety (or any other emotion)." In these studies, S's autonomic responses were measured and shown to vary as a result of M's stressful experience and as S and M shared FSA's. One difficulty in utilizing the empathy concept, Stotland admitted, was that merely degree of arousal could be measured by GSR, the quality of the emotion could not be. Thus, if M received shocks, S's heightened arousal could be interpreted either as empathy or as sadistic joy.

Strength of Identification and Vicarious Processes

In the present formulation, identification is distinguished from empathy, imitation, and other vicarious processes. Identification is conceived as a quality of the relationship between people, a type of bond between people, which shall be explicated more fully later.

Identification is viewed as an intra-personal and interpersonal <u>structure</u>, analogous to the anatomical structure of an artery. Empathy, on the other hand, is an observable <u>event</u> or behavior, one typically measured by autonomic activity. Imitation, too, is a behavior, generally a more discernably overt one than is empathy.

Vicarious processes, such as empathy, imitation, vicarious classical conditioning and extinction, are all considered events, but events requiring a structure along which to proceed. To complete the analogy, these processes are similar to blood circulation, which takes place within the arterial structure. In empathy, to speak figuratively, the affect flows across people along the bond that is identification.

Identification, unlike an artery, however, is a structure that is not visible. Its existence, and its value as an intervening variable, must be inferred from other observable events. In particular, we shall claim that findings about variations in the intensity of vicarious experience can be better ordered by invoking the concept of identification. Relationships which conform most to the condition of high identification, that is, where the bond is strongest, are those most conducive to the transmission of intense vicarious processes. The data that are available to test this view tend to confirm it. In the Bandura and Huston (1961) experiment, for example, a relationship was allowed to develop between M and S over three experimental sessions. For half the Ss, M acted nurturantly; we can assume that the relationship quickly became a reasonably strong union. For the remaining Ss, M acted aloof; the relationship here was no doubt a stunted one. S's imitation of M's novel responses was pronounced only for the former, and was particularly marked for those

Ss who had been independently assessed as "dependent." These latter Ss no doubt generally tended to develop strong bonds with others.

Most of the Stotland studies can be viewed in this way also. Whenever FSA's were experimentally established, introjection was in evidence. Lack of FSA's resulted in but little empathy or similar (vicarious) self-appraisal. The existence of FSA's are connotative of strength of identification, for here the S "sees himself" in M, surely one of the meanings of the identification bond.

The evidence becomes more compelling, however, when the bond develops naturally, in the real world, rather than being produced in the laboratory. In these cases, it is easier for us to believe that identification truly exists in a strong form. Kagan and Phillips (1964), for example, had children observe a contest between one of the parents and a stranger. The children smiled more frequently when it was announced that their parent succeeded than following the stranger's victory over the parent. Moreover, cardiac acceleration was greater when the parent was victorious.

A study by Brenner (in preparation) is particularly compelling. In previous work, he required Ss, who were seated around a table, to flip over a card, in their turn, and read aloud the neutral word which was written upon it. After two cycles (i.e., after each S had read two words) they wrote down as many words as they could recall. These were plotted as a function of the seating distances from

S to the speaker of the word. Brenner found that Ss recalled words read by others several positions away better than those whose positions were within two adjacent spaces on either side. He interprets this result as the interfering effects of arousal which precede and follow performance, even so simple a performance as this. In the critical study, he utilized couples who had a several month old dating relationship. In one condition, for example, only the girls in each of the couples read the words, and at the conclusion of the reading, the males were asked to list those words they could recall. It was found that the farther the reader of the word sat from his girlfriend, the better was he able to remember the reader's word. Brenner reasons that, as in the original experiment, the girlfriend experiences arousal when her performance is impending or just completed which impedes her recall of contiguous words. This arousal gets transmitted to her observing boyfriend, he gets aroused <u>for her</u>, which impedes <u>his</u> recall in identical fashion. Only when M was a girlfriend or boyfriend, however, did the vicarious effect materialize. Moreover, it was most pronounced for those couples which were observed by E to be touching each other. Any study which shows, as this one does, that any kind of arousal or heightened drive state is experienced by one who identifies with the person undergoing the arousing treatment, but not (or to a lesser degree) by one who does not so identify tends to confirm

the hypothesis. In the Berger (1962) study, for example, if both Ss who did and did not identify with M had been used, no doubt the former would show more striking vicarious classical conditioning than the latter.

A Reconceptualization of Identification

Many writers, we have seen, use either imitation of M, or an affective tie with M, or both, as the defining condition of identification. We have argued that imitation is only useful when it applies to the identification of children, but not that of adults--who do, surely identify. For the sake of parsimony, we asserted, the identification concept should ideally be equally applicable to all instances--hence we proposed the term imitation be confined to a separate, though related, phenomenon.

Here, the issue of S's affective tie with M will be examined. The difficulty with employing this as a defining condition is that it is too broad, too general. An affective tie means, in the final analysis, anything other than a neutral evaluation of another. There are many people who I like--or I hate, for that matter--but with whom I do not identify. To use the existence of an affective tie as the defining condition, however, would disallow this distinction. We must conclude that it is not a simple affective tie, but something more restrictive that is intended by the term identification.

It is not necessary, similarly, that S see a

similarity between himself and M. Such a distinction is a troublesome one, because it requires the specification of a degree of perceived similarity, such that short of this point identification is said not to exist, and beyond it always to exist. Furthermore, it requires the selection of dimensions of similarity which are and are not crucial. Stotland seems to have eluded these problems by accepting any amount or dimension of perceived similarity as sufficient for inferring the existence of identification, but actually he has encountered the pitfall of making his definition nearly universal, for nearly everyone can see some similarity between themselves and any given other.

Finally we wish our conception not to hinge on S's admiration or attraction to M, or to M's goal-states. If we take defensive identification seriously, we must be prepared to accept instances where S hates or is repulsed by M but still identifies with him. Moreover, S must not be said to identify with all whom he admires, nor with all whose goal-states he wishes he possessed. I admired Hitler's power and wealth but did not identify with him.

While we reject all the above considerations (imitative behavior, real or perceived similarity to M, attraction to M, affective tie with M, as well as desire to command M's resources) as defining characteristics of identification, we acknowledge that they can (sometimes with great regularity) be associated characteristics. We reject them only as the feature which distinguishes

exemplars from non-examplars of identification.

Instead we ask whether S has made a more or less conscious decision to "invest" a part of himself in M, to commit to M's care a portion of himself. This commitment need not require M's knowledge or concurrence. It requires that S chooses to intermingle his fate with that of M. He does this by incorporation of M and his relationship to M into his self-picture. In part, S comes to define himself in terms of M. S identifies with M if and only if S's self-definition concerns M.

It is useful to relate this investment concept to the concept of sub-identity developed by Miller (1963). By sub-identity Miller meant salient divisions of the total identity. This notion is appealing in the present context. I identify with another, if this relationship forms a sub-identity for me. Typical sub-identities, for Miller, include son, husband, father, professor of psychology, etc. If any of these are the actual sub-identities for a given person, then the related others, i.e., father, wife, other psychology faculty, etc., are others with whom he identifies. In this approach, the concepts of identification and identity are brought into close theoretical proximity where they belong, but where other theorists have been reluctant to put them.

Kuhn and McPartland (1954) have devised a method which enables a good view of the concepts of sub-identity and identification, which they call the 20-question Who

Am I test. The S is required to write twenty different answers to the question "Who am I?" The responses to this are an indication of the sub-identities and the others with whom S identifies. (It is not a particularly apt measure of identification, however, for many reasons. For example, S desires to help E define or categorize him, instead of or as well as defining himself for himself, which confounds the issues.) Walker (1965), using this measure, found that Blacks (in a Black southern school) used a racial term about twice as often in defining themselves as did whites in a white school. Whites, however, gave nationality responses (e.g., American) about three times as often as Blacks did. This implies that Blacks tend to identify with other Blacks and whites with other Americans.

When a person identifies with an other, that is, defines himself in part in terms of this other, he is giving to that other a responsibility, an investment; M becomes liable for a portion of S's self-definition. Generally, it is hypothesized that such an investment requires some trust in the other, which normally develops only after some satisfying interaction. Thus, according to the present conceptualization, identifications which E attempts to create anew in the laboratory are rarely of the force or quality of real-life identifications. This argument, however, does not imply that there is no identification with the stranger, simply a minimal one, for

whenever S has the sub-identity "man," he identifies with all men. No single strange man, however, carries much of S's total investment.

Investigation of the variable of strength of identification is, generally speaking, methodologically treacherous. It requires either manipulating different levels, or measuring it, or both. We have argued that actual manipulations are difficult or uncertain. A wiser course is to select subject populations which can be assumed a priori to differ on this dimension. A useful technique, for example, is that used by Kagan and Phillips (1964) and Brenner (in preparation). They chose S-M pairs with naturally occurring identification bonds (child-parent, and sweetheart, respectively), and varied strength of identification by comparing these to the vicarious processes of S to not-parent or not-sweetheart M's.

Measurement of strength of identification is also frequently desirable or necessary. The measures in the literature, as summarized by Bronfenbrenner (1958) are diverse, usually methodologically unsound, and based on what we have argued are unwise defining characteristics. Yet to measure commitment to or investment in M, our preferred definition, seems a formidable task as well. Nevertheless, in the present study, an attempt is made to measure this attribute, which we term "cognitive identification."

In the case of some instances of identification

bonds, one or more of the defining characteristics which we rejected might be suitable as well. The point is that only investment in M, definition of self at least in part in terms of M, is suitable across all kinds of bonds, father-son, son-father, national and racial identification, etc. "Affective Identification," the degree to which S is attracted to, likes, or evaluates positively, M, for example, might be an appropriate measure of racial identification, or sports fan identification, but would be inappropriate where identification with the aggressor were a possibility, as in son-father, athlete-coach, or employer-employee identification. Given the admitted difficulty of measuring cognitive identification, if appropriate given the kind of identification, it might be prudent to take measures of affective identification, as well. To the degree that defensive identification is ruled out, the two measures should agree with one another, and relate to other variables in identical fashion.

It can be seen that what we have called the identification effect, that S's self-appraisal will be enhanced following M's successes, and deflated following M's failure, becomes a plausible concomitant of cognitive identification. We have posited that when S identifies with M, M's activities are influential with regard to S's self-appraisal. If S defines himself in terms of M, i.e., if cognitive identification is strong, M's successes and failures will be S's, i.e., the identification effect will follow.

Interpretations of the Identification Hypothesis

The identification hypothesis, then, may be summarized as Zander (1958) states it, "The characteristics of a person may rub off on one who is identified with him, but not on one who is not identified with him . . . a person's evaluation of himself can be affected by the successes and failures of those with whom he is identified." (pg. 108).

In this section, we will distinguish five explanations of the identification hypothesis, as well as review the evidence regarding its tenability.

The first interpretation argues about the difficulty of drawing self boundaries. Is my body self, or not self? As William James (1890) puts it:

> The empirical self . . . is all that he is tempted to call by the name of me. But it is clear that between what a man calls me, and what he simply calls mine, the line is difficult to draw . . . In its widest possible sense, however, a man's self is the sum of all that he CAN call his, not only his body and his psychic powers, but his clothes and his house, his wife and his children, his ancestors, and his friends, his reputation and work, his lands and horses and yacht and bank account . . . If they wax and prosper, he feels triumphant; if they dwindle and die away, he feels cast down. (Vol. 1, pg. 291)

French (1963), delivers a similar argument using the term identity where James uses self. Thus, those with whom I identify might all be considered in an important sense, a part of myself; their accomplishments are merely a special case of my own. We shall call this

interpretation "extension of self."

A second interpretation of the phenomenon, call it the "achieving membership interpretation," argues that in some circumstances, my association with this other may be considered itself an accomplishment for me. Marriage, and membership in select groups, have this feature. If I have been chosen by this other, the greater his accomplishments, the greater my own in having been worthy of this association. In instances where the association between self and other was thrust upon the two, such as having been born into a nationality or race, this interpretation would predict no self-reappraisal following other's performance, for there is no reflection on self if other was not "achieved."

The third possibility, termed here the "How I'd do" interpretation, holds that as M and S are more similar, the M's performance will be interpreted by S as a good estimate of how he himself will perform. Many of Stotland's already reviewed experiments confirm this interpretation. Indeed, the FSA$\longrightarrow$DSA theory is a formal statement of it. In the Burnstein, Stotland and Zander (1960) experiment, for example, if S was told he was similar to M, he assumed he would have like ability as a deep sea diver. A study by Shupp (reported in Diggory, 1966) also obtained a finding supporting this interpretation. S predicted his later performance as conforming to the performance of whichever of two M's his previous performance

was most nearly like. A later study by Cutick (1962) replicated this finding. M was described as similar or dissimilar to S, on dimensions separate from the ability tested. Only when M was similar, did S's prediction of his later performance match the feedback he was given regarding M's. In both the latter two studies, the effect only obtained when M succeeded, not when he failed. Similarly, Raven and Fishbein (1965) and Fishbein, Raven and Hunter (1963) asked Ss to predict what scores they would obtain on an exam. Being told the scores of others had significant effects on these predictions. What distinguishes this explanation from those previous is its ability-specific nature: the self-appraisal is a self-evaluation. If self-esteem is affected, it is so only <u>via</u> a more specific evaluation. A husband's business success, this interpretation would have to argue, enhances his wife's self-esteem, <u>because</u> she predicts she could be a success at business, too; <u>because</u> she introjects his business acumen and calls it her own. This is not to say that the "how I'd do" process does not operate; simply that it is not a sufficient explanation of all instances of the identification hypothesis. It is likely, for example, that the self-esteem of the rabid alumnus would increase after his school's football victory, yet surely it is not because he reevaluates his own football ability.

A fourth position holds that the identification hypothesis follows from symbolic interaction considerations.

If S is identified with M by others, and M succeeds, these others reassess their estimation of S; S then "reads" their reassessment and reappraises himself. The wife of the business executive is accorded more respect by others the more he is successful, since she is associated with him in their eyes. This respect by others, according to the usual Meadian hypothesis, soon gets translated into self-respect. According to this reasoning, it is not who S chooses to identify with that is important, but with whom others see him identified with. This hypothesis could not account for any occurrence of the identification process in which S was not seen by others as associated with M, as in, for example, secret hero worship with celebrities. Fortunately for this interpretation, no such data exists, though it is likely, for example, that many who identified with the late President Kennedy experienced self-enhancement over his handling of the Cuban missile confrontation, though their identification with him was a private affair.

The final explanation, termed "belong to good team," posits that identification means that S sees himself and M as a team competing against other such teams. The M's success is taken as evidence of the "home team's" superiority. Simply being a member of a good team has beneficial effects on self-esteem. This explanation is probably merely an elaboration of the first, the extension of self interpretation. When a representative of the team

succeeds or fails, S's estimation of the worth of the team waxes or wanes. S, being a member of this team, shares in his own assessment. Even Mead (1934), the proponent of another set of principles, concedes this one: "Self-evaluation may to some extent be determined by the superiority of the group of which one is a member" (pg. 208). Rosenberg (1965) presents some evidence that people of higher socioeconomic status, and of more approved racial characteristics, have higher self-esteem. Izard (1959) replicates the former finding. Finally, Brown (1965) who, in his section, "Interdependence of the Self-Conception and the Conceptions of Others," adopts a stance similar to the one portrayed here with regard to the conception of identification, writes:

> Conceptions of groups, even geographically remote groups, can affect the self. Consider the following rather common transformation in the thinking of a Negro who has grown up in the United States in recent years. We may suppose that when he was a boy, African Negroes seemed to be the painted savages who chased a heroic white Tarzan in Saturday afternoon movies, and American Negroes the forever inferior descendents of these savages, and himself an average American Negro. When he grows older, he learns that: (1) the art of African Negroes is as highly regarded by some intelligent people as is the art of Europeans; (2) African colonies can become African nations and produce their own leaders who are treated deferentially by European and American heads of state; (3) American Negroes, the descendants of Africans, sometimes go to college, enter professions, become famous artists or statesmen. All of this has implications for his self-concept. He comes to think of himself as a young person of some ability and good character who has been given no opportunities.

> Treatment by whites that once seemed reasonable enough and certainly inevitable now seems a wrong that must be righted . . . A change in a man's impression of certain geographically remote others--Africans--might in this way effect a transformation in his impression of himself and those nearest him. (pg. 651-652)

Summary

The question impelling this investigation is whether an individual's appraisal of self can be affected by knowledge of the performance of others. The literature on identification with others was reviewed, since it was argued that if S identifies with another, that other's performance will have similar implications for the self-appraisal of both the other and S. This we termed the identification hypothesis. The contributions of various theorists of identification were reviewed and evaluated, which led to the reconceptualization of the identification concept, a reconceptualization from which the identification hypothesis became a plausible consequence. The evidence for this hypothesis and several plausible interpretations of it were then critically examined.

It was also mentioned early in the paper that the literature on social comparison processes had propositions to offer with regard to the central question of how knowledge of the performance of others affects self-appraisal. This literature, and its ramifications for the central question will be examined in the remaining section of this chapter.

Social Comparison Processes

Festinger's theoretical paper on social comparison processes (1954) addressed the problem of not only assessment of abilities but of the correctness of one's attitudes as well. Indeed, Schachter (1965), and others, as Latane (1966) notes, have extended the work further, to cover assessment of emotions and bodily states, as well as that of personality characteristics. However, in the present review we shall confine ourselves to research on only that assessment relevant for our own purposes, namely, appraisal of abilities.

Festinger's first assumption is that there exists in the human organism a drive to evaluate his abilities. Uncertainty about one's abilities generates attempts to obtain more precise information. How this information is obtained, and what are the consequences of this search, form the basis of the theory. Physical, objective data are preferred, he asserts, but social data are desirable also, and often the only kind available. Moreover, the objective data may not, in themselves, be sufficient to evaluate the ability: "One might find out how many seconds it takes him to run a certain distance, but what does this mean with respect to his ability--is it adequate or not." (pg. 119)

In short, Festinger concludes, people compare their abilities with others in order to evaluate themselves.

Festinger draws heavily on the previous literature on level of aspiration (LOA) in this part of the formulation, equating the person's stated LOA with the person's evaluation of his ability. This seems somewhat wrong-headed, since ability evaluation seems better approximated by the score S thinks he will get, than the score he will try for. However, since the former, termed level of expectation, is merely slightly exceeded by the latter, the LOA, and since the two seem to be governed by much the same principles (Lewin, et al., 1944), the error is not a serious one.

Previous work had shown that performance relative to LOA is akin to subjective feelings of success and failure. Thus, a person whose performance exceeds his LOA reports feeling the thrill of success, while if he falls short of LOA, he reports that he is experiencing subjective failure. Festinger noted that a series of studies (Gardner, 1939; Festinger, 1942; Chapman and Volkmann, 1939; Hilgard, Sait and Magaret, 1940; Hansche and Gilchrest, 1956) found that LOA, and by implication, feelings of success and failure, could be made to vary depending upon the S's standing in a social group. The most well-known of these studies, the Chapman and Volkmann (1939) study, is illustrative. College student Ss were given the score supposedly obtained by one of the following groups: literary critics, other students, or WPA workers. The score given was always the same. Ss

set their LOA at a level below that obtained by literary critics, and above that of WPA workers. Moreover, as Anderson and Brandt (1939), and others have shown, when others like himself have scores different from his own, his LOA always moves closer to the level of performance of others.

Thus, Festinger concluded, people compare their abilities to that of others in order to evaluate themselves. An important point is that the more similar the others, the more precise the evaluation. Hence, if given a choice, Ss will prefer to compare with similar rather than dissimilar others. This similarity could be either on the ability in question, or in other respects. A major deduction from this assertion is that LOA's are stable when the person is performing close to the group mean. Since stable evaluations (LOA's) are the goal of comparison, it is argued, attempts will be made to perform close to the group-mean--preferably by bringing the group mean close to self, rather than vice versa.

Several studies tested the above proposition. Radloff (1966) found that non-average Ss had lower correlations of true and estimated scores for a trial--i.e., less accurate evaluation--than did average scorers. Moreover, the range of LOA given by S over a series of trials was smallest for average scorers. However, if many similarly deviant others could be compared with, he found, stable self-evaluation could be reestablished. Dreyer (1954)

obtained a similarly supportive result: a smaller proportion of Ss who had been given average feedback than those told they were non-average (superior or inferior) reported their ability as "fair." Dreyer interprets this result as indicating that non-average performers have unstable self-evaluations, and hence prefer to place themselves in a catch-all category.

Brehm and Festinger (1957) and Hoffman, et. al., (1954) tested the proposition that if given the opportunity, Ss would affect the scores of others in such a way as to narrow the range of scores, i.e., improving the performance of low scorers, hindering the performance of high scorers, so that their own self-evaluations would be more stable. When Ss were given some power to control the scores of others, when the task was important, and when the others were similar to them, this turned out to be the case.

Wheeler (1966), Hakmiller (1966) and Thornton and Arrowood (1966) addressed the issue of who S will choose to compare with when given the choice. In each of these experiments, the S was one of several "others" receiving scores; S always got the median score and was allowed to see the raw score of any other rank. Festinger (1954) hypothesized that, because there is a unidirectional drive upward for abilities, a drive to do better and better, S will prefer to compare with someone else of slightly superior ability. Wheeler obtained this expected result,

especially under conditions of high motivation. Hakmiller, however, found that if the individual receives unexpectedly negative feedback, he chooses to compare with the most inferior possible other, as if to preserve his self-esteem by using the comparison by which he would remain in the most favorable possible light. If the comparison is valuational, as well as evaluational, argues Hakmiller, other motives besides accurate assessment become operative. The Thornton and Arrowood study confirms this result.

Festinger, Torrey and Willerman (1954) contended that as the group becomes more attractive, the greater will be the importance attached to differences in ability, and hence larger differences in self-evaluation will result. In their experiment, after the manipulation of attraction, three group members were given similar scores and the fourth member was given a somewhat lower score. The results showed that for low scoring Ss, lower self-evaluations were made by those high than by those low in attraction for the group, while for the higher scoring Ss, higher self-evaluations were made if the member had high than if he had low attraction for the group.

Gerard (1961) and Jones and Gerard (1967) make the distinction between reflected and direct appraisal; the latter is the sense Festinger uses the term, while the former concerns what S perceives the others think of him given his score. Reflected appraisal should magnify the

direct appraisal effects the more S is reliant on the group. Gerard (1961) confirms this expectation.

As Singer (1966) notes, the comparison can be expected to affect not only self-evaluation, but self-esteem:

> People do not compare with others in order to evaluate only one opinion or ability. Implicitly, they are also evaluating their opinion of themselves. In the general case, they are evaluating their self-esteem. When a person asks, "how much X do I have?" he is also asking, "what sort of person am I for possessing that much X?" (pg. 105)

We can assert, then, that social comparison--knowing the performance scores of others--affects self-appraisal in both its specific and general sense. Exactly what is the effect? Singer (1966) writes: "It is not sufficient for a person to know how much X he has; he must also know the distribution of X." (pg. 107). As the distribution of X gets higher, the person suffers a diminution of self-appraisal. As the distribution gets lower, the person enhances his self-appraisal. The better the performance of one to whom I compare myself, the worse I look by comparison; I look better the more inferior the performance of the other.

This process, which we shall call the social comparison effect, does not operate for the performance of all others, however. It depends, as Pettigrew (1967) notes, on the similarity of the other, in his terms, only if the other is a reference person. He observes that Blacks, for

example, expect whites to perform better, generally, and do not often use the performance of whites for comparison purposes, though Lefcourt and Ladwig (1965) show that in some situations, they do. And Davis (1966), in a compelling use of the theory, finds that students tend to evaluate their academic abilities only in comparison with their fellows on the same campus, not students nationally.

It is a failure of the theory, however, that it cannot easily predict in advance whether a given other will or will not be considered by S as a reference other.

Two experiments demonstrate quite neatly the operation of the social comparison effect. Gergen and Morse (1967) had potential job applicants sit next to a student who played the role of either Mr. Clean or Mr. Dirty. Mr. Clean was immaculately dressed, carried an attache case with a slide rule and several sharpened pencils, and acted self-assured, while Mr. Dirty was inappropriately dressed, confused, and carried a dog-eared copy of The Carpetbaggers. S's self-esteem changed upward significantly after confrontation with Mr. Dirty and downward after confrontation with Mr. Clean.

In a pilot study by Braver (1964), Ss were asked to help validate a self-esteem measure by responding as honestly as possible. Allegedly to "make sure we get a representative sample," Ss were first asked to respond to some demographic items, namely grade-point average, college board scores, father's income, and number of

extracurricular activities they regularly took part in. These responses were made on a sheet, which, supposedly had been completed by all previous subjects. For half the Ss, the "previous subjects" were generally high on all these attributes, for half they were low. Though the results did not achieve significance, in large part because of the small N's employed, they were strongly suggestive of the fact that Ss in the "superior others" group subsequently reported lower self-esteem than Ss in the "inferior others" group.

Thus, the social comparison effect has been empirically demonstrated. If another is a referent other, (a) the better his performance, the lower S's self -esteem will become and (b) the worse his performance, the higher, S's self-esteem will rise.

CHAPTER 2

SYNTHESIS AND DEVELOPMENT OF HYPOTHESES

In the previous chapter, in order to examine the question of how information about another's performance can affect an observer's self-appraisal, we reviewed two main bodies of literature, that on identification and that on social comparison processes. From the former we derived the hypothesis that, when S identifies with another (or group of others), the other's success will be met with increases in S's self-appraisal; while the other's failures will be met with decreases in S's self-appraisal. From the latter we derive a second hypothesis: when S uses another (or group of others) as a reference, that other's success will be met by a decrease in S's self-appraisal, while the other's failures will be met with increases in S's self-appraisal.

The reader has no doubt noticed two points concerning these hypotheses. First, they specify different assumptions--the one an identificatory other, the other a referent other. However, very often the person with whom there is identification also serves as a referent person. Genuine teams, for example, generally serve this dual purpose. One both identifies with one's teammates and

uses their performance as a bench mark against which to evaluate his own. There are a wide range of situations where the activating conditions of both hypotheses are simultaneously met.

Which brings us to the second point: the two hypotheses predict opposite outcomes. When another who is both an identificatory and a referent other succeeds, the identification effect predicts he will raise, while the social comparison effect predicts he will lower his self-appraisal. When the other fails, the identification hypothesis predicts S will lower, the social comparison hypothesis that he will raise his self-appraisal. We have a situation where two principles, each straightforwardedly derived from a large body of literature which contains substantial empirical support, are unavoidably in conflict.

Is the conflict real, or merely theoretical and academic? Consider some examples: (1) a man whose wife performs outstandingly well; (2) a typical ghetto Black who learns that other Negroes are beginning to achieve; (3) a bench-warmer on the New York Mets when observing his teammates spectactularly winning the Pennant. In each of these cases, the tendency for S to identify is countered by his tendency to compare himself. When a husband succeeds in business, our intuition would predict his wife should undergo self-enhancement; yet clinicians have daily examples where she feels more inadequate the greater his accomplishments.

In everyday terms, whenever there is the opportunity to be proud of another for his achievements, there is also the opportunity to be envious of him. We must conclude that the theoretical conflict accurately reflects one which often actually occurs in real life. This fact has been noticed by Heider (1958) in his chapter, "Reactions to the Lot of the Other Person:"

> If we take into account only the value dimensions of an experience, i.e., its positive or negative character, four types of reactions to the lot of the other person may be . . . distinguished: (1) that o has a positive experience is positive for p; (2) that o has a negative experience is negative for p; (3) that o has a positive experience is negative for p; (4) that o has a negative experience is positive for p. In the first two types, the relation between p and o is syntonic or concordant. It presupposes a sympathetic "identification;" the first may be called sympathetic enjoyment, the second simply sympathy or compassion. The last two types show a discordance or antagonism between p and o; the third may be called envy, the fourth malicious joy or Schadenfreude. (pg. 277-278)

Berger (1962) prefers the term empathy for the first two types, which we call, simply, the identification hypothesis, and sadism for the fourth. In our terms, (3) and (4) are both instances of social comparison effects.

Recognizing, then, that a conflict exists, two courses are open. We can attempt to uplift one theory and attack the other, or we can acknowledge the virtue in both positions and proceed to attempt to clarify, both conceptually and empirically, the conditions under which each

principle takes predominance. Considering the bulk of supportive data on both sides, and our acknowledgement that real world interaction presents the same conflict (to identify or to compare) as theoretically exists, the second course is the wiser to follow. We shall, then, take as a paradigmatic situation one in which S can both identify with the others and/or use them for a referent. We shall observe the effect on S's self-appraisal as a function of the feedback he is given about performance of the others. We shall, additionally, argue for a series of concomitant conditions, some situational, and some concerning S's personality characteristics, which should favor one or the other principle in resolving the conflict. The arguments will be of such a nature that their test against empirical data can be permitted.

Our hypotheses, then, are hypotheses about <u>conditions</u> which favor identification effects and those that favor social comparison effects. We shall consider identification effects to have been found when (a) S enhances his self-appraisal following M's success or (b) lowers it following M's failures. Social comparison effects are found when (a) S lowers his self-appraisal following M's success or (b) enhances his self-appraisal following M's failures. Two measures of self-appraisal are used: self-esteem is the primary measure, with the more specific ability evaluation a supplemental measure. In addition,

several other indices of ego-protection, related to self-appraisal, will be employed. The above statements summarize the goals of the present research.

Situational Variables

Others' Success vs. Failure

Although both the identification and the social comparison propositions include predictions of both the situations where the other fails and where he succeeds, it is possible that one or both operate only unidirectionally. It is plausible, for example, that people are predisposed, generally, toward upward changes of self-esteem. This would imply that if M succeeded, S would tend to identify, whereas if he failed, social comparison would be more likely. It can be seen that if this occurred to very great extent, however, (for example, if it overwhelmed the other variables) the concepts of identification and social comparison would become superfluous; the single principle of a tendency towards self-enhancement would explain the results more parsimoniously. It is also possible that while social comparison operates in both directions, identification effects may operate only in an instrumental sense, only to enhance the self--only M's successes may alter S's self-appraisal, while his failures are disregarded. Such an argument is made by Kagan (1958) who reasons that, in the case of M's failure, the

identificatory bond is weakened, but no introjection takes place. Shupp (reported in Diggory, 1966), Cutick (1962) and Zander, Stotland, Wolfe (1960), all obtained this unidirectional result. As our conception of identification differs both operationally (they used M's who were formerly strangers to S) and theoretically from that of the investigator's above, we tentatively hypothesize instead, that either principle can operate in both the case where M succeeds and where he fails, depending upon other variables. We do, however, permit a test of this hypothesis, by including both conditions in the experimental design. The success and failure of other, which we refer to as "Team Success-Failure" becomes the first situational variable in the factorial design.

S's Knowledge of Own Performance

Implicit or explicit in much of the writing on social comparison is the notion that S knows his own score on the ability-test which he compares with the scores of others. Recall Singer's (1966) phrase, "It is not sufficient for a person to know how much X he has; he must also know the distribution of X." (pg. 107). Knowledge of the distribution affects his <u>evaluation</u> of the score. A score looks good if it is on the high end of a distribution of the scores of referent others, but bad if it is on the low end. Two data points are necessary for change in evaluation by this account: S must

know S's score in addition to the score of the other(s).

Other research purportedly concerning social comparison does not, however, provide S with his own score. The Raven and Fishbein (1965), Fishbein, Raven and Hunter (1963), Festinger (1942), and Chapman and Volkmann (1939) studies, for example, give S the scores of a category of others and asks him to predict his own. Generally, these are categories dissimilar to his own, but often a condition of a similar other is included. Most interest is directed to the effects on S's ability-evaluation or LOA after this data--particularly the data for groups obviously his inferior or superior--is given. It is important to note, however, that essentially this same paradigm has been used by researchers interested in identification (Shupp, 1966; Cutick, 1962; the Stotland program)--though their interest is primarily in the effects on S's ability-evaluation after data for <u>similar</u> others is given.

The general issues of whether or not a score for S is necessary for social comparison--and whether without this score the process is one of identification or social comparison remain a problem. Let us address the second question first. The data in studies of investigators of both persuasions are similar: if the social categories for which a score is given are generally inferior or superior to S's, he believes he will perform better or worse, respectively. If the category is S's own, he estimates he will do similarly. It is not the data but the

interpretation which is at issue. It would be in error to dispense simply with the issue by claiming that when other is dissimilar social comparison results, and when he is similar identification effects result. This would contradict a fundamental proposition of social comparison, namely, that it takes place primarily with similar others.

Rather, the most appealing answer to the controversy for the case when S has very little independent means to guess his score is to call the process neither one of social comparison nor identification. The reason for this is that both processes, as we have discussed them, result in S's reevaluation of himself. In the situation in question, however, i.e., when S has no independent data about his own performance, solely information about that of others, no reevaluation of self takes place. Instead, S evaluates only the task; he establishes a frame of reference for how various social categories, including his own, will do. No subjective experience of success or failures follows from the input of information--he changes only his standards for performance. When there is a strong identificatory bond, to be sure, S might experience vicariously the subjective feeling of success or failure--as in the sense of the "belong to good team" interpretation discussed in Chapter 1. This should occur primarily when it is clear that S's social category is competing with another. It is incorrect, however, according to this view, to call S's setting of LOA identification effect

when there is no evidence that he vicariously experiences success or failure, that he merely reevaluates a task, rather than himself.

This brings us back to the question of whether social comparison can occur when S gets no score for himself. According to the above argument, it cannot, for S does not reevaluate himself or experience success or failure. The argument, however, was carefully articulated to apply only to the case where S has no external information about his standing on the attribute to compare with another's standing. The fact that he merely has no objectively given score does not necessarily mean that he has perfect uncertainty about where he stands. Often, related abilities or previous information about the present ability can allow him to make an educated guess regarding his score. Whenever S is reasonably confident about such a guess, the activating conditions for social comparison processes are met, and reevaluation can take place. The more confident he is in this guess, the more dramatic should be the effects of social comparison. We can assume that receiving actual objective feedback about his score, defines the pole of this certainty dimension.

What of Ss who do not receive this feedback, who do not have this degree of certainty? If S has participated in the task but not received knowledge of his results, it would be an error to assume that here S resides in perfect uncertainty about his standing on the ability.

"Personal feedback," proprioceptive and other cues regarding the difficulty S encounters in performance no doubt give him some indication about his standings. It is, however, impossible to say with certainty whether S can, on this basis, be sufficiently confident about his estimate to engage in actual social comparison, i.e., self-reappraisal, when he learns the scores of others. We do predict, however, that Ss with full knowledge of their own performance are more prone to engage in social comparison than those without full knowledge. Since social comparison means suffering decline in self-appraisal as others succeed and increases when they fail, this hypothesis amounts to a hypothesis that the more fully S knows of his own performance, the more dramatically his self-appraisal will raise when M fails and drop when M succeeds.

Operationally, Ss in the Score condition will take the test and receive a score for themselves as well as for Ms, the others. Those in the No Score condition have incomplete knowledge of own performance; they always take the test but do not receive a score for themselves, only for the Ms. (The score Score Ss receive always stays constant, whatever the scores given the others, that is, regardless of whether in the Team Success or Team Failure condition.) The experimental prediction corresponding to the above hypothesis is that a two-way interaction, Score-No Score (we'll call this variable, simply, Score) by

Team Success-Failure, on the dependent variables of self-appraisal will be found.

The Nature of Competition

In Chapter 1, several possible interpretations of the identification hypothesis were enumerated. The final one, termed the "Belong to Good Team" interpretation, was preferred because it seemed to be applicable to all hypothesized instances of the identification hypothesis, while the others had more limited explanatory potential. It can be seen that "Belong to Good Team" can actually be reduced to a social comparison effect, if it is assumed that when one identifies with another, the two are seen as a team which competes against other such teams. When a husband succeeds, his wife interprets this as a victory for her family over other families, much as she would if she herself succeeded. Social comparison for individuals works as a zero-sum game competition: if they win, I lose, and vice versa. For teams, it operates analogously. What is crucial here is the conception of self and H as a team or unit.

It follows that as this team conception is prevented, the identification effect will not appear. In fact, if a team-member can be induced to compete with other members of his team, the traditional social comparison effect should ensue. Depending upon the situation, a person can see himself either as an ally or a competitor or rival to another. We hypothesize that

when he sees himself as an ally, identification effects will predominate, but if he construes the other as his rival, the social comparison effect will emerge.

Operationally, we will instruct subjects in the Within Team Competition condition that they are competing against other members of their same team, to "see which does best;" while those in the Between Team Competition condition will compete as a Team, against a rival Team. The above hypothesis corresponds to an experimental prediction that social comparison effects for the Within Team group, and identification effects for the Between Team group, will be found. Since identification effects signify upward changes in self-appraisal in the Team Success and downward changes in the Team Failure condition, while social comparison effects are the reverse, this amounts to a prediction of a two-way interaction, Competition by Team Success-Failure, on the dependent variables of self-appraisal.

There is a large accumulation of literature with regard to self-evaluation as a function of either cooperation or competition with team members. Johnston (1966, 1967; Johnston and Nawrocki, 1966, 1967) finds that when feedback is given to the team as a whole, S is likely to blame his teammates for its failure, but share with them credit for success. Raven and Rietsema (1957) found confirmation of the fact that as cooperation or alliance with team members becomes more salient, identification

effects, "empathizing with the group experience," increased. In the Gergen and Morse (1967) study, competition with the other failed to increase social comparison processes more than non-competition, but a manipulation check revealed that the variable was not well manipulated. Perhaps the most relevant study is one by Ferguson (1958). She noted that siblings can either be in competition or in alliance with one another in gaining status in the family, and that particular sib-pairs are predisposed toward one or the other of these orientations. On the basis of pre-experimental interviews, she divided her sib Ss into these two categories (in the present study we manipulate, rather than select, to obtain differences on this independent variable). She obtained a social comparison effect for the competitive pairs and an identification effect for the allied pairs when they evaluated themselves after participation in an experimental task.

Situational Variable Summary

In Table 1 is schematically represented the hypothesized effects. The hypotheses regarding situational variables will be confirmed if both a two-way Score by Team Success-Failure interaction and a two-way Competition by Team Success-Failure interaction regarding self-reappraisal are found. In Table 2 is presented the design.

TABLE 1

A SUMMARY OF THE HYPOTHESIZED EFFECTS OF THE PROCESSES AND EXPERIMENTAL CONDITIONS ON REAPPRAISAL OF SELF. ARROWS REPRESENT THE PREDICTED DIRECTION OF REAPPRAISAL.

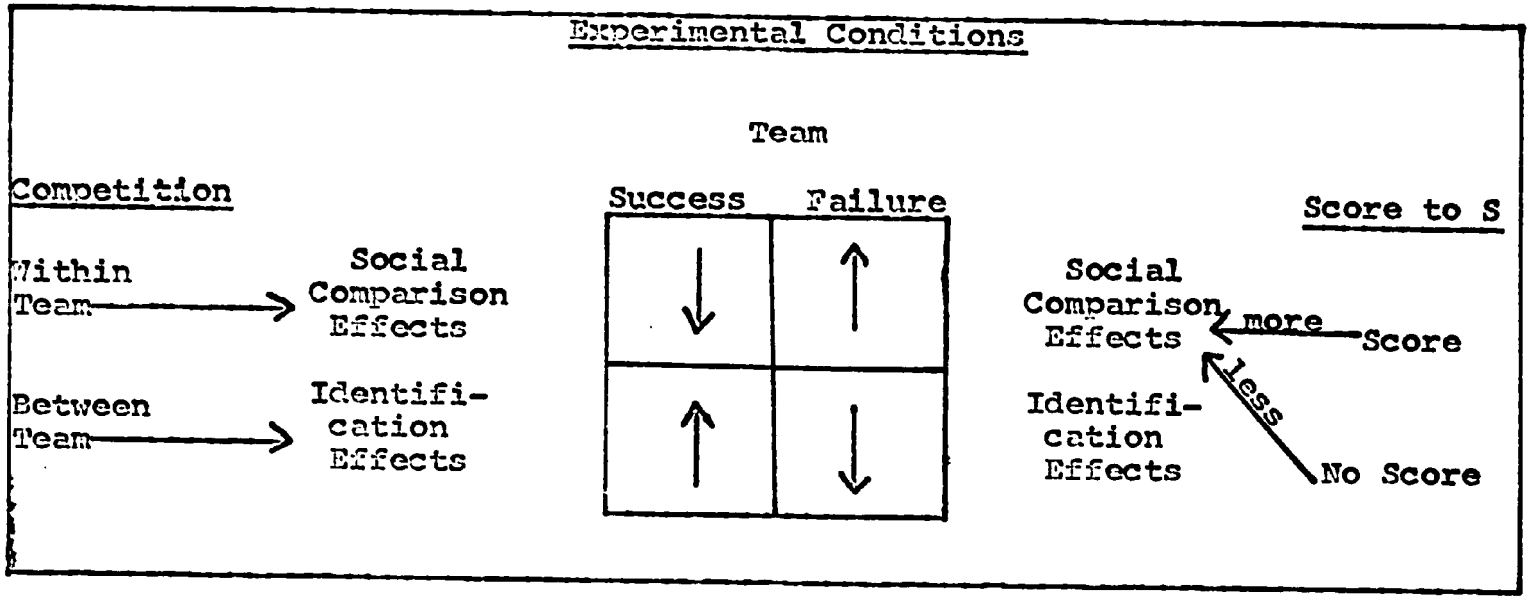


TABLE 2

A SUMMARY OF THE 2 X 2 X 2 FACTORIAL CONDITIONS.

	Competition			
	Between Team		Within Team	
	Score	No Score	Score	No Score
Team Success				
Team Failure				

Personality Variables

Though we manipulated situational variables to provide for their experimental test, personality variables, of course, are given. Two techniques are possible: (1) we can select from our population groups which fall at different points in the distribution and insert them factorially into the design with other treatments; or (2) we can simply do post hoc correlational analyses on the natural distribution. The latter course has been adopted.

Strength of Identification

Though we selected groups in which natural identification is assumed to be strong, it is unavoidable that

individuals will vary with regard to the strength of identification. Hence, we measure it and develop hypotheses concerning its interaction with other variables in affecting identification processes and social comparison processes.

In Chapter 1 it was argued that as strength of identification increases, so should vicarious processes, in this case, identification effects in changes in self-appraisal. The issue, however, is not so simple, for strength of identification with another also often means (and in the case of the present type of identification probably does mean) attraction to other. As Festinger, Torrey, and Willerman (1964) demonstrated, attraction increases social comparison processes. The Score-Within Team Competition condition captures almost exactly the paradigm used by Festinger, et. al. (Though they attempted to manipulate attraction to group, their manipulation failed. Their results are based instead upon a correlation-like technique, as in the present study.) Our hypothesis, then, must be dual. We predict greater identification effects, the greater the strength of identification, in those situational conditions which favor identification. We predict greater social comparison effects, the greater the strength of identification, in those situational conditions which favor social comparison. It summarizes these two hypotheses to say: changes in self-appraisal in directions hypothesized for situational

variables will be higher as strength of identification is higher.

This is tantamount to a prediction that an analysis of covariance on the criterion variable of self-reappraisal, with strength of identification as the variable covaried out, will yield greater effects for the components of variance originally predicted to be significant than would a simple ANOVA. We wish, then, to test the gains of ANCOVA over ANOVA, the most straightforward implication of this hypothesis. Unfortunately, there exists no accepted significance test of such an hypothesis.

We instead look at the cells in the design where the effects are hypothesized to be most pronounced. Both Within Team Competition and Score, we hypothesize, favor social comparison effects. Both Between Team Competition and No Score are most conducive to identification effects. Strength of identification is hypothesized to increase identification effects in the former and increase social comparison effects in the latter; identification effects mean increases in self-reappraisal following M's successes, decreases following his failures, while social comparison effects imply the reverse. This amounts to prediction that there will be positive correlations of strength of identification, as measured both by "cognitive identification" and "affective identification," with (upward) self-reappraisal in the Between Team Competition, No Score, Team Success Condition and its complement, the Within Team

Competition, Score, Team Failure condition. There should be negative correlations in the Within Team, Score, Team Success condition, and its complement, the Between Team, No Score, Team Failure condition.

Initial Level of Self-Esteem

Many writers (e.g., Kagan, Stotland) have argued that those low in self-esteem are those most prone to identify. This position has been explicitly taken, for example, by Kagan (1958) and by Sanford (1955), who argued that low self-esteem is the reason one identifies. Presumably, since the identification process can be used in an instrumental sense, to vicariously enhance self-esteem through the accomplishments of M, those most likely to resort to it are those whose self-esteem can stand enhancement. Those of a more self-confident demeanor can rely more on themselves for self-enhancing accomplishments.

Research on the question, primarily in the Stotland program, has had mixed results. Some studies (Stotland and Dunn, 1962; Stotland and Hillmer, 1962) report confirming data, some (Cutick, 1962) inconclusive or marginal findings, while other experiments have found high self-esteem Ss more likely than lows to identify (Stotland and Cottrell, 1961, 1962).

A second hypothesis can be added with regard to the relationship between self-esteem and self-reappraisal.

It is conceivable that initial level of self-esteem operates in a self-perpetuating way. Stotland, et. al., (1957), for example, report that:

> Persons high in self-esteem . . . are characterized by tendencies to protect themselves from negative self-evaluations . . . persons with high self-esteem . . . evaluate an objective failure as a small failure and an objective success as a large success; persons with low self-esteem . . . evaluate an objective failure as a large failure and an objective success as a small success. (pg. 56)

In the present No Score condition, it is possible that high self-esteem subjects estimate, in the absence of contradictory information, that they have done better than their referents; while low self-esteem subjects assume they have done more poorly; in both cases, social comparison processes will be activated in the No Score condition--where we have argued they are least likely.

We tentatively formulate two hypotheses: (1) that low self-esteem subjects are more likely to identify than high self-esteem subjects. Furthermore, we expect that (2) given the same feedback, high self-esteem subjects are likely to interpret it favorably, low self-esteem subjects to interpret it unfavorably. These amount to the following predictions: (1) We predict those low in self-esteem will most strongly produce identification effects which means increases in self-appraisal in the Team Success condition and decreases in self-appraisal in the Team Failure condition. Thus we predict a negative correlation

of initial level of self-esteem with (upward) changes in self-esteem in the Team Success condition, and a positive correlation in the Team Failure condition. In addition, (2) we expect, over all conditions, the correlation between self-esteem and self-reappraisal (change in self-esteem and ability evaluation) to be positive. With regard to change in self-esteem this implies that the positive correlation in the Team Failure condition is predicted to be more positive than is the negative correlation in the Team Success condition to be negative.

Field-Dependence

Several other individual variables have been hypothesized to relate to the proclivity to identify. For example, Stotland and Dunn (1962) have argued that authoritarianism and birth order distinguish persons with regard to this predisposition. Another possible individual difference variable is field-dependence. As Witkin, et. al., (1962) conceive it, people differ in their perceptual and cognitive style, one major component of which is their tendency toward psychological differentiation. People who are field-independent achieve great differentiation of objects in their environment, including themselves. Those who are field-dependent encounter difficulties, conceptually as well as perceptually, in making distinctions, in differentiating. As they argue, this difference in cognitive style extends

to affect one's picture of himself as an individual entity. Field-dependents "have a less developed sense of their identity and of their separateness from others" (pg. 5), while perceiving in a field-independent way contributes to a "person's . . . awareness of his own . . . attributes . . . as distinct from the . . . attributes of others . . . internal frames of reference have been formed and are available as guides for definition of the self." (pg. 134).

We hypothesize, accordingly, that field-dependents will be more prone toward identification effects than independents, while the latter will exhibit greater social comparison effects. This hypothesis corresponds to the following prediction: We predict those high in Field Dependence (high scores means Field Independent, low scores mean Field-Dependent) to most strongly produce identification effects. They should, then, enhance their self-appraisal when M succeeds , diminish it when M fails. Thus, there should be a negative correlation of Field-dependence with (upward) self-reappraisal in the Team Success condition, and a positive correlation in the Team Failure condition.

Summary of Hypotheses & Experimental Predictions

(1) A two-way interaction Score by Team Success-Failure on the dependent variables of self-appraisal will reflect our hypothesis that giving S an explicit score

increases his tendency to engage in social comparison processes.

(2) A two-way interaction Competition by Team Success-Failure on the dependent variables of self-appraisal will test our hypothesis that when the other is a rival social comparison effects will predominate, but when the other is an ally identification effects are the stronger.

(3) As strength of identification increases, the magnitude of the above hypothesized effects should increase. Correlations of strength of identification, as measured both by cognitive and affective identification, with change in self-appraisal will be positive in the Between Team, Score, Team Failure condition, and the Within Team, No Score, Team Failure condition; they will be negative in the Within Team, Score, Team Success condition and the Between Team, No Score, Team Failure condition.

(4) Those whose initial level of self-esteem is low should demonstrate more visible identification effects. We predict a negative correlation of initial level of self-esteem with (upward) self-reappraisal in the Team Success condition, and a positive correlation in the Team Failure condition.

(5) Field dependent Ss should exhibit greater identification effects than field-independent Ss. This means there should be a negative correlation of Field Dependence (a low score means greater field dependence)

with self-reappraisal in the Team Success and a positive correlation in the Team Failure condition.

(6) Initial level of self-esteem acts to color new feedback in such a fashion as to maintain self-esteem at the present level. We predict, for the sample taken as a whole, a positive correlation (rather than the negative one predicted by regression toward the mean) between initial level of self-esteem and self-reappraisal.

CHAPTER 3

METHOD

Subjects

The choice of subjects for this study represented an even more crucial decision than is usually the case for social-psychological investigations. This is because of the complex interpersonal nature of identification. We have theorized that an identificatory bond is a tenuous one, one that grows up slowly in the course of interaction. This theoretical viewpoint precludes attempting to create identification anew in the laboratory, by, for example, using a set of instructions similar to those used by Back (1951). Instead we must investigate cases where an identificatory bond already exists in a natural group or dyad. The difficulty with this solution, methodologically speaking, is that there are several types of groupings (i.e., bonds) that fulfill our basic definition, and to whom we would like to generalize our results; for example, husband-wife, parent-child, identification with team, racial identification, national identification, etc. At the same time, because of the logistics of experimental research, only one can be selected for study in a given

experiment. But there are enough unique features to each of these types that we cannot know whether the findings utilizing one type of grouping would have been obtained had another type of identificatory bond been utilized. This problem would be minimized if we had allowed ourselves to think that the bond between two strangers that we could create in a laboratory setting was equivalent to that found in the natural world, for then we could generalize more easily. Our contention that "real" identification is fundamentally different from anything we could simulate by instructions has required us to choose an instance of the former and be satisfied with some inability to generalize until further research with different types of subjects and bonds confirms earlier conclusions.

The type of identification chosen for this study was sex-role identification. It is a commonplace observation that young boys have strong tendencies to adopt masculine behaviors, and this manifests itself, among other ways, by strong in-group feelings with other boys, and rejection of girls and anything feminine. In other words, young boys identify with other boys, in the sense that they invest in other boys, that they define themselves in terms of the attribute "boyness." They use the category boys as a positive reference group, while girls are a negative reference group.

In our culture, however, girls probably have weaker sex-role identification, a fact borne out by research

evidence concerning masculinity-femininity (See Lynn, 1969, for a summary of this literature which suggests that boys are more masculine than girls are feminine.) Rejection of male enterprises is substantially weakened in girls and often, as is the case with tomboys, boys form a positive reference group for girls. Thus, since strong identification was a prerequisite criterion for subject selection, the population was restricted to young males.

This strong sexual identification in males, however, declines with age, and in most cases burgeoning interest in the opposite sex at puberty nullifies the earlier allegiances. It was necessary, therefore, to choose the age of the subjects wisely. We wished to choose boys who were young enough to have intact identificatory bonds with other boys, yet old enough to be able to read, comprehend and follow instructions and manipulations, and complete the various measures in a meaningful enough fashion so that a successful test of the hypotheses was assured. It was decided that boys of age 10-11 best satisfied these dual requirements.

Accordingly, the subjects were fifth and sixth grade boys from two elementary schools in a Detroit suburb. Ten different classrooms, with ten different teachers, were utilized.

Demographic Characteristics

As the suburb was an exclusively working class area, the boys were all of lower socio-economic class. In this school district, whose classification of "academically depressed" had recently been removed, approximately 15% of the children went on to college. There were problems of lack of motivation in academic pursuits, and also of intellectual inability in the students. According to the principal, only about 30% of the present sample performed on standardized national examinations at their own grade level or above, while fully 20% tested two or more grades below the grade to which they were assigned. In one case, a sixth grade subject (whose data, naturally, was deleted from all analyses) was completely unable to read, and his teacher read him all the measures aloud. In several classes, approximately half of the students were classified as "educable retardates". It is obvious, then, that the attempts to choose an age-group which jointly maximizes strength of sex-role identification and verbal facilities were at least slightly frustrated on the second criterion because the sample was a scholastically below average one. A total of 134 males were tested. In addition, 105 females in these classes provided data which was eventually discarded.

Design Overview

The study was divided into two sessions. In Session I, E entered the classroom. Students whose parents had responded to a permission slip (See Appendix A.1) that they did not wish their child to participate in the study were excused and/or given other work to occupy them. (Roughly two students per average class of 25 were lost in this way.) The remainder, both boys and girls, first completed identificatory information, an initial self-esteem inventory, a strength of identification and field-dependence measure. Then they were given instructions on how to take the bogus experimental test and completed it. When these were collected, E left the room, answering any questions about "getting the tests back" nonconmittally and instructing the classroom teacher to do the same (See Appendix A.2). This session lasted about an hour. The teacher assisted in answering questions about word meanings and instructions, but otherwise was dissociated from E.

Session II took place from four to seven days later. Only boys took part in this session. Randomly formed groups of three to five boys, all classmates and only boys who had participated (successfully[1]) in Session I were led into another room in the school building. If

---

[1]Two Ss failed to complete the measures fully and correctly. These boys did not participate in Session II.

the group had been assigned to one of the experimental conditions, they were given instructions stressing the importance of the bogus test and what it tested and the independence of the tested ability from other scholastic abilities. Then they were given feedback on the test appropriate to the condition to which they had been assigned. Immediately afterward, they completed the post-self-esteem inventory and a supplemental questionnaire. Finally they were dehoaxed and dismissed, after cautioning them not to relate the events of the session to their classmates. For experimental groups, Session II lasted about one-half an hour.

The control subjects, also in groups of three, four, or five, merely completed the post-self-esteem inventory upon entering the laboratory. No further instructions or feedback were given this group. These procedures took about ten minutes per group.

The classroom teachers were given instructions to answer any questions that arose between sessions in a noncommittal way. (See Appendix A.2).

Instruments

Session I

Self-esteem. The primary dependent variable, self-esteem, was measured by the Coopersmith Self-Esteem Inventory (SEI) (Coopersmith, 1968). This instrument was

designed for use with an 8-12 year-old sample. In its original form, it consisted of 50 items all keyed to self-esteem. For each item the subject was to respond "LIKE ME" if the statement described how he usually felt, and "UNLIKE ME" otherwise. Sample items are: "I am pretty sure of myself"; "I would rather play with children younger than me." Complete reliability and validity data are given in Coopersmith, but suffice it to say here that the scale performs more than satisfactorily on these criteria. In fact, one issue that was contended with in utilizing it in the present study was its extraordinarily high test-retest coefficient of .88, with as long as a six-week between test interval. If the scale measured an attribute that had this degree of stability, perhaps no experimental manipulation could produce changes. However, since Gergen and Morse (1967) had obtained predicted changes using an adaptation of this scale (although with a college student sample), it was decided that the attribute measured by the SEI scale was sufficiently malleable to justify its adoption. Since we wish in the present research to measure change in self-esteem, and direction of change, not merely relative magnitude, was of considerable importance, i.e., since we wished to have both a before and an after measure of self-esteem, the SEI was divided into two halves, one administered at Session I, the second immediately after the feedback manipulation of Session II. The division into halves was made in the following way:

Coopersmith, after computing an item analysis, had constructed a "SEI Short Form" consisting of the 25 items which had the highest item score versus whole score reliability coefficient. The items that appeared on this half can thus be considered better indices of self-esteem than those that did not. We wished an equal number of "good" (and "bad") items to appear on each half we constructed. Also, pre-testing with the present sample had shown that response bias, particularly acquiescence response set, might cause some distortion in results. The SEI is not composed of an equal number of positively keyed and negatively keyed items. If the keying, in addition to the content, of the items makes a difference, it is clearly desirable that the two halves each possess the same number of positively keyed (and thus negatively keyed) items.

Accordingly, a 2 x 2 table was constructed with direction of keying (positive versus negative) one dimension and goodness of item reliability (appears on short form versus does not) the other. The fifty items were classified into the four-fold table. The resulting table is presented in Table 3.

TABLE 3

CLASSIFICATION OF THE 50 ITEMS ON THE SEI (COOPERSMITH, 1968) ACCORDING TO THE DIRECTION OF KEYING AND THE GOODNESS OF ITEM RELIABILITY.

		Direction of Keying		
		Positive	Negative	Total
Goodness of item Reliability	Appears on SEI Short Form (Good)	8	17	25
	Does not appear on SEI Short Form (Bad)	10	15	25
	Total	18	32	50

From each of the two cells with odd numbers as entries, one item was eliminated, one at random, the other because girls must answer it differently than boys. Finally, a random procedure was followed within each cell such that half the items were assigned to the Session I pretest version, the other half to the Session II post-test version of the measure. The resulting halves are randomly parallel forms in a true sense. The half administered as a before measure in Session I is presented in Appendix A.3, together with the written and oral instructions to subjects in completing this form.

Identification. There were two measures of identification used in the present research, one which attempted

to measure an attribute termed "cognitive identification" and the second measuring "affective identification". By cognitive identification was meant the extent to which the subject invested in boys, or defined himself in terms of other boys, that is in terms of the attribute "boyness". By affective identification was meant the degree to which one approved of, liked, and evaluated positively, boys as opposed to girls. As stated earlier, it was thought that empirically, in many cases of identification, the two types were highly correlated, that one normally liked those that he identified with; but that the concepts were certainly conceptually distinct, and occasionally empirically independent, as in the case of identification with the aggressor.

It was felt that the concept of cognitive identification implied the items by which it was measured. If S defined himself in terms of the attribute of "boyness", the opposite of the attribute "girlness", it followed that he saw a lot of differences between boys and girls, but little differences among boys. Furthermore, it implied that S saw himself as a representative exemplar of his team; and finally it implied that he felt team-members' (boys') activities should be restricted, as his was, to commerce with his own sex. An item was drawn up to capture each of these implications of the concept. The items, with their written and spoken instructions on how to respond, are given in Appendix A.4.

To measure affective identification, the positivity of the evaluation they gave to boys, more standard techniques were available for adaptation to the present study. The measure consisted of ten person-descriptive adjectives which were chosen from the list by Anderson (1964). The words chosen were extreme in evaluative rating, five being positive and five negative words, and were selected from the list according to the criteria of extremity, familiarity of meaning (and ease of reading) to a sample of this age, and the applicability of the adjective to members of either sex.[2]

The adjectives used, and their evaluative ratings, as given by Anderson (1964), on a 7-point scale where 0 equaled very unlikable and 6 equaled very likable, are given in Table 4.

---

[2]Pre-tests had shown that two adjectives were consistently attributed only to members of one sex, "mean" to boys, "helpless" to girls. This resulted in decreasing the variability (and hence the reliability) of the scale. These two adjectives were replaced in the final version reported here.

TABLE 4

THE TRAIT-ADJECTIVES USED ON THE AFFECTIVE IDENTIFICATION INSTRUMENT AND THEIR EVALUATIVE RATINGS, AS GIVEN BY ANDERSON (1964)

Good Traits		Bad Traits	
Adjective	Rating	Adjective	Rating
honest	5.55	lazy	1.26
friendly	5.19	grouchy	1.17
interesting	5.11	bossy	1.12
smart	4.85	nosey	1.02
active	4.55	selfish	.82
Mean	5.05	Mean	1.08

The S was asked to decide whether each adjective better described girls or boys. The score on this measure was computed by adding the number of positive words atributed to boys to the number of negative words attributed to girls. The entire scale, and the oral and written instructions for its completion, are given in Appendix A.5.

Field Dependence. There are a number of rather standard techniques that have been used to measure field dependence. Witkin's (1954) first measures were behavioral and involved complex apparatus, such as the Rod and Frame and Body Adjustment Test. Later, however, he adapted

Gottschaldt's Hidden Figure Test by adding lines and colors, etc. The result was the Embedded Figures Test (EFT), a speed test which takes about an hour to complete. Goodenough and Eagle (1963) developed a children's version (the CEFT) which was appropriate for five-to nine-year olds. Kaufmann (1965), however, correlated the original Gottschaldt test with the EFT on a group of children and found Part A of the former, a 10-minute component, to correlate .67 with the latter in 8-year olds, and .51 with 15-year olds.

The measure used in the present study, then, was the original Gottschaldt (Thurstone, 1944) Part A. The instructions were lengthened and adapted and four sample problems were given. The examiner worked the samples on the blackboard and helped Ss individually when they found the task incomprehensible. For Part I, five minutes were allowed; two and one-half minutes were provided for Part II. The number of correct recognitions of the sample figure within the complex figure constituted the score; for Ss who had finished (either part of) the test within the time limit, their score was extrapolated to yield the score they could have achieved in the time limit had there been more items, by solving for X in the formula:

$$\frac{\text{Number correct}}{\text{Time they took}} = \frac{X}{7.5=\text{total time allowed for test}}$$

If a S finished within the time limit, he was instructed to raise his hand, and the time, to the nearest quarter minute,

was recorded. The test and instructions appear in Appendix A.6.

Experimental Test. A bogus task was designed especially for this experiment. The purpose of the task was to have something about which credible feedback regarding performance could be given. Thus, the bogus task was designed with the following criteria in mind: 1) that it be easily administerable to a group; 2) that there be little personal feedback, that is, that Ss learn little about their success by observing their own performance; 3) that the tested ability be different enough from skills and abilities about which the S already had knowledge so that past performance could not serve as a guide in predicting his success; and finally, 4), that the tested ability favor neither boys nor girls. With regard to the last three criteria, it was of paramount importance that personal or team outcome could not be predicted by past experience, or else the experimental feedback would either be incredible, or of little information value. To fulfill these aims, a test labelled the Sizes and Shapes Test was designed and is presented in Appendix A.7. It consisted of five parts, each of four items, two of which were especially designed for this measure and the remaining three were adapted from standard optical illusions. A question at the conclusion of this test asked how many items the S estimated he got wrong on

the 20-item test and served as a check on the degree to which the test met its objectives.

Session II

The instrument used in measuring self-esteem in this session is presented in Appendix A.8. Both control and experimental groups completed the second half of the SEI. In addition, experimental groups answered eight extra questions, some of which were manipulation checks, others of which were supplementary measures of ego-protection, related to self-appraisal. These questions are presented in Appendix A.9.

Manipulation of Conditions - Session II

Ss in the control group were simply told to complete the second part of the SEI, of which "the form might look familiar, but the questions are different from those you answered before."

The procedure for experimental Ss was to first stress the importance of the bogus task. This was done by claiming that the Sizes and Shapes Test was a measure of how well they "sized up new situations", alleged to be an important ability. The specific instructions were:

> You boys might remember that a few days ago you were given quite a few tests. One of them looked like this (hold up), and it was called Sizes and Shapes Test. It was the one where there were 20 problems where you answered A or B for each one. Do you remember this? This was a very important test, you know, because

> it shows how well you are able to size up new situations and arrive at their correct solutions. People who do well on this test usually size up new situations correctly, so they have a good understanding of the things they have to do in order to get along well. We have given this test many times, and we usually find that if a child does well on this test, he will be able to get along very well as he grows older. Children who don't do so well on this test often have a harder time and may not get things done exactly the way that they want. That is why this test is so important.

Next, an attempt was made to dissociate the ability from other abilities about which the S may have had secure self-evaluations.

> Also this test is not like arithmetic or reading tests you have taken. You may do very well on this test even if you usually don't do too well in reading or arithmetic or other school subjects - or you may usually do well on school subjects, but when it comes to the Sizes and Shapes Test you might not do well. In other words, you can't judge how well you will do on this test by how well you do on other tests you may have taken in school before.

At this point, the experimental groups were divided along the dimension of competition. Half of the groups were told:

> Now we have tested several classes of fifth and sixth graders in other schools, so we know how well fifth and sixth graders can do on this test and how many they usually get wrong. In some of the classes, the boys we tested did better than the girls; in other classes, the girls seemed to size up new situations better than the boys. We want to know what will happen in <u>your</u> class. We want to know whether the girls in this class do better on the test than the boys, or do the boys do better than the girls.

This created the Between Team Competition condition. The remaining groups were told:

> In each class, some of the boys we test usually do better on the test than other boys. Some boys seem to size up new situations well, others not as well. We want to know which boys will be the ones to do well on the test in your class. We want to see which boys do well on the test, and which boys do less well.

Ss given this instruction are in the Within Team Competition group.

The feedback of team scores was then administered. Every S in the experimental room at one time, i.e., in the group, got identical feedback about how the Team, boys, performed. Only boys in distinct groups got different feedback with regard to whether the Team succeeded or failed. By contrast, whether or not the S got a score for himself varied within the group. Half of the Ss in each group were given a score for themselves, while the other half were not.[3] The instructions read:

> In your class, we haven't had time to grade all of the children's tests to see how well everyone did on it. We have only graded eight people, all boys so far, so we only know how well about half of the boys in your class did. I am going to hand out all these scores for you to see. You will see that there are eight scores on your sheet. If you were one of the boys whose test we graded so that we know how well you did, your score is circled. If the sheet I give you has a circle on it, the circled score is your score.

---

[3] In the case of groups with an odd number of members, the odd person was assigned to the condition with the smallest n to that point.

> If your sheet does not have a circled score on it, that means that your test was not graded, so we don't know how well you did; and because we ran out of time, we won't be grading any more boys' tests.[4] Now, I am going to hand you back these sheets, together with some other forms that have questions we would like you to answer. When I hand you these papers, you may look at the scores for a while. It is very important though that you don't talk to each other right now.

At this point the feedback sheets were passed out. For the Team Failure condition, there was one score of nine wrong, two of 11 wrong, and four or five scores of 10 wrong. For the Team Success condition, there was one score of three wrong, two of one wrong, and four or five scores of two wrong. For the Score condition, a score of six wrong was substituted for one of the predominant scores, and this score was circled. In other words, feedback for the No-Score condition consisted of either eight scores of 9, 10, or 11 wrong, or of 1, 2, or 3 wrong, while for the Score condition, there were only seven scores on these extremes plus a circled score of six wrong. Thus, Ss in the Score condition received a score that was either better than that of their poor performance classmates or worse than that of their successful classmates, but always the same score. All Ss learned that their classmates

---

[4]The inclusion of this phrase, it was hoped, lessened the S's natural preoccupation with his <u>own</u> performance, and decreased the salience of his guessed score (which could very possibly act in identical fashion to objective score) so that the No Score condition would be as different as possible from the Score condition.

performed quite homogeneously, either well or poorly, and if there was deviation, it was their own.[5]

Next the following instructions were read. For the Team Success condition: "As you can see, almost all the scores are one, two or three wrong. Scores that are one, two, or three wrong are really quite good, and the boys in this class who got these scores seem really to have done well on this test, and are very good at sizing up new situations." Added for the Between Team Competition condition was the phrase: "In fact, these scores are so high that I would be surprised if the girls in the class will be able to have such high scores, too. So, even though I won't know for sure until I grade the tests for the girls I'll bet the boys in the class will do better than the girls." For the Team Failure condition the instructions were: "As you can see, almost all of the scores are 9, 10, or 11 wrong. Scores that are 9, 10, or 11 wrong are really not too high, and the boys in this class who got these scores seem not to have done too well on this test and are not very good at sizing up new situations." Added for the

---

[5] A possible alternative arrangement would have given No Score Ss an uncircled score of six, thus providing that all Ss within the Team Success or the Team Failure condition see the identical set of scores. This solution, however, would have had two undesirable consequences: 1) It would have compromised the monolithic quality of the others' scores, thereby perhaps weakening both identification effects and social comparison effects. 2) Score Ss would learn that the others were monolithic, while No Score Ss would learn that they were not, a case of non-equivalence on a possibly confounding variable.

Between Team Competition condition was the statement: "In fact, these scores are so low, that I would be surprised if the girls in this class aren't able to do better than this. So, even though I won't know for sure until I grade the tests of the girls, I'll bet that the girls in this class will do better than the boys." All subjects were then instructed: "Now turn the page and begin to answer the questions, making sure to put your name, school, and class at the top. The form might look familiar, but the questions are different than those that you answered before. When you have finished this page, turn it and start the next. If you have any questions, raise your hand. Do not talk to the other children."

The Ss then completed the post-SEI and supplementary questions.

Dehoaxing

In many social psychological experiments, especially where potentially damaging self-information and/or deception is employed, dehoaxing involves carefully explaining the falsehood of what they were earlier told, the reason for the deception, and the true state of affairs. With the present sample, however, a better procedure was to give new information which was uniformly favorable and which could supersede the earlier feedback. Accordingly, after collection of the previous measures, which marked the end of the experiment proper, the following

instructions were read:

> One trouble with the Sizes and Shapes Test (hold up) that I didn't mention before is that a lot of people who really do size up new situations very well, and who will get along very well as they get older, don't do as well as they should have the first time they take the test. Sometimes they were just too tired, or not feeling well the first day, or maybe they just improve a lot with practice, or maybe because I am a stranger to you I somehow caused you not to do as well as you would have if your regular teacher had given you the test. We don't know, but we have found that if the person takes the test a second time, often he does much better, and shows that he really is able to size up new situations. So I am going to give you the test over again right now. It's how well you do on the test this time that really counts and that will show how well you really size up new situations. When you have finished the test, bring it right up to me and I will grade it right now, so we can know how well you really size up new situations.

(E passes out tests and pretends to grade them as they come in.)

> Well, you boys have all done very well this time. Each of you really had a very good score, and this is the time that really counts. So I see each one of you here really sizes up new situations very well for a fifth or sixth grader. I thought that all the boys in this class could really do well on this test, and you did: Congratulations.

This procedure appeared to be extremely successful.

Finally, they were told that:

> I would like to ask you boys not to talk to the other kids in class about what happens here. I would like all of you to keep my secret about what happens here, at least for a few days. Is that O.K. with you? Will you keep it a secret? If you think you can

> keep what happens here a secret, raise your hand. Thank you. I'll take you back to class now.

After this, the group was escorted from the experimental room to their classroom, and a new group of 4 or 5 called, until all the potential Ss in the class had been utilized.

Randomization Procedures

Both order of administration of conditions, and assignment of Ss to conditions, were controlled by randomization according to the following procedures. Since there were four different types of experimental groups (it will be recalled that the Score variable varied within a group), and since we wished the N for the control group to be twice as large as that for any experimental group, random permutations of the numbers 1-15 were obtained, and using Mod 5 arithmetic, the order of experimental groups was dictated.

The names of successful[6] Session I Ss were then permuted randomly, and upon entering a classroom for Session II, the first several names were read, absent Ss being, of course, disregarded, and so the desired size group was formed. The size of the group was determined by the total number of potential Ss in the class, and by

---

[6]See footnote 1, this chapter.

an attempt to keep the cells even. The first and third names in each group were always the ones given Scores, while the second and fourth were always in the No Score condition; in cases where a five man group was formed, the placement of the fifth S into either the Score or No Score condition was determined by an attempt to keep the growth of N equal in the various cells.

CHAPTER IV

RESULTS--AND THEIR INTERPRETATION

This chapter is organized as follows: The first section reports data bearing on the adequacy or effectiveness of the experiment in providing a bonafide test of the hypotheses. We shall be concerned there with the reliability of the measures; the degree to which the instructions were successful in creating differences between conditions, on the independent variables; and the success of the randomization in equating groups with regard to the variables measured in Session I. Following that will be the section reporting tests of hypotheses involving the situational variables. The final section reports the effects of the individual difference variables on the dependent ones.

Adequacy of Measures, Manipulations, and Randomization

Measures

There are two criteria of the adequacy of psychological instruments that psychologists are most concerned with: reliability, and validity. The former is an index of the degree of internal consistency, the extent to which

the instrument measures some one attribute. Validity tests question whether that attribute measured is, in fact, the one the scale was intended to measure, a question answered, in general, by asking whether the variable associates with others as it theoretically should.

This section reports only reliability, i.e., internal consistency, information. The primary index will be the coefficient α, the soundest and most recommended computation available, or its adaption to the case of dichotomous responses, the Kuder-Richardson Formula 20 (KR-20). (See Nunnally, 1967, for a detailed discussion and derivation of these formulations.) Validity information will be given in the form of tests of hypotheses and other relationships, in later sections of the chapter.

SEI. Since Coopersmith's original scale was adapted and halved for use in this study, it is necessary to inquire about the adequacy of even so standardized a scale. The KR-20 for the Session I half was .716 (N = 134; k = 24). For the Session II half, the KR-20 was .771 for all 120 Ss. However, since in Session II, prior to measurement by the SEI, Ss underwent treatments intended to influence self-esteem, perhaps to fragmentize it, it is necessary to inquire whether the various experimental groups differed with respect to the KR-20 for the Session II SEI. Table 5 presents these data. These coefficients are all closely enough bunched that they can

be considered no more than random deviations about a common value.

TABLE 5

KR-20 COEFFICIENTS OF THE SESSION II SEI FOR THE SEVERAL TREATMENT CONDITIONS

Condition	Between Team		Within Team	
	Score	No-Score	Score	No-Score
Team Success KR-20	.506	.699	.801	.483
N	12	12	12	12
Team Failure KR-20	.797	.775	.822	.821
N	12	12	12	12
Control		KR-20	.853	
		N	24	
All Experimental		KR-20	.730	
		N	96	
All Ss		KR-20	.771	
		N	120	

Often test-retest reliability is taken as a measure of internal consistency, but this has the additional component of changes in true score over testings. Thus, generally, test-retest r should be lower, somewhat, than the α which is, according to classical test theory, a measure of the correlation which would be obtained between two tests, at one administration, whose items are chosen

at random from a common infinite item pool. As the procedures here were largely in conformity with the above underlying model, the inequality, test-retest r less than α was expected to hold. As a comparison of the data in Table 6, which displays the r's, to Table 5 demonstrates, this result was obtained.

TABLE 6

CORRELATION COEFFICIENTS OF SESSION I (PRE-) TO SESSION II (POST-) SEI SCORES, BY CONDITION

Condition		Between Team		Within Team	
		Score	No-Score	Score	No-Score
Team Success	r	.202	.599	.827	.724
	N	12	12	12	12
Team Failure	r	.672	.748	.558	.928
	N	12	12	12	12
Control	r		.728		
	N		24		
All Experimental	r		.650		
	N		96		
All Ss	r		.667		
	N		120		

The test-retest r's are substantial and not significantly different across groups (Test given in Winer, 1962, p. 599; $F = .99$; $d.f. = 7,80$; $p > .10$). In addition, the various

treatment combinations which intervened between the two halves did not change the slope of the regression line.

The fact that the test-retest r's are very nearly as large as the α's is in agreement with Coopersmith's (1967) findings that test-retest correlations were abnormally high. (The disparity in correlations, the present .667 versus his .88 can be attributed almost entirely to two factors: (1) The smaller number of items in the present version, a rather mechanical factor in decreasing reliability; (2) the fact that the items were different on the two halves here though identical in the Coopersmith data. If S remembered how he responded to specific items and wishes to appear consistent, the agreement between his scores on first and second administrations would be increased. This, of course, could not occur if the items were different.)

Cognitive Identification. The coefficient α for this scale was .13. This was unacceptable. In addition to the small number of items, there was relatively little commonality in what was being measured by each of them. When the worst item, Question 3, which correlated negatively with the others, was dropped, α jumped up to .31, which however, was still unacceptable. Hence, the cognitive identification measure was left out of all further analyses.

Affective Identification. This ten-item scale had a KR-20 value of .74 which was quite respectable. As predicted, it correlated substantially with the remaining three items of the Cognitive Identification scale. As a result of this correlation the reliability (α) of a combined index of identification, adding together the ten-item scale with questions 1, 2, and 4 of the Cognitive Identification measure, became a robust .95. However, no later analysis using the affective measure is changed by more than a few decimal points if the combined index were substituted. Hence, for conceptual clarity, only the affective measure will be implied hereafter when reference is made to "the strength of identification".

Field-Dependence. This measure is quite standard, and results in previous research justify its inclusion in the present study. In addition, as a speed-test rather than a power-test, conventional techniques (e.g. α KR-20) are inappropriate as estimates of internal consistency. Thus, no reliability data is presented here.

Bogus Task. No score was computed for this task, and no real ability involved, hence no reliability data was taken. However, the task was designed with specific desiderata in mind. One of these regarded the degree to which the S would be in doubt about his performance. An indication of this is afforded by the magnitude of the variance of the scores Ss guessed they got immediately

after completion of the task. The range of possible scores was 0-20; S.D. of guessed scores was 3.59 (N = 134). Six S.D.'s thus cover more than the entire range. The dispersion of scores is thus large. The large dispersion implies that little information consistent across all Ss was transmitted by merely taking the task, without the addition of objective feedback. A second indication of the lack of certainty Ss had about their performance is given by another statistic. Of those Ss given No Score, if there were a high correlation between the score they guessed they got before, with the score they guessed they got after receiving feedback about how their classmates did, it would suggest that the latter input was minor compared to personal feedback. If the correlation were low, though, their guess of their own score would likely have been greatly affected by information of how others did; personal feedback would probably be low. The r = .20, which is insignificant (d.f. = 47).

A second datum of interest regarding the bogus task is the mean of the guessed scores. Recall that Ss in the Score condition were told that their own score was six wrong. If this score, arbitrarily selected, was radically different from the mean predicted score, the average S would, in addition to any other information conveyed by his score, be learning that he did far better or worse than he expected. This could have had confounding effects on the dependent variable of interest.

Fortunately, the mean predicted score was 5.5 (N = 134), very close indeed to the score given. Considering only those Ss in the Score condition, who are most affected, the mean was 5.54 (N = 48).

Supplementary Questions. Immediately after the Session II SEI was completed, Ss responded to a number of supplemental questions. Of these questions, Question 2 and 6, "How good are you at sizing up new situations?" and "How satisfied are you with your ability to size up new situations?" were intended as two items measuring the same attribute, namely, evaluation of their ability (A-Eval) in sizing up new situations. A score for each question was obtained by assigning a 5 to the positive end and a 1 to the negative end of the scale. The reliability of this two-item scale was .606. Scores were obtained for the other questions in the same way. Question 7 and 8, S's liking for the boys in their class who were, and were not, their "good friends", respectively did not correlate positively (r = 0.078, N = 96), so they are, in later analyses, treated separately. Similarly, when the responses to questions 4 and 5, "How well did the other boys in your class do on the Sizes and Shapes Test?" and "How good are the other boys in your class in sizing up new situations?" are added together, a scale with α = .567 results. This scale,

a measure of perceived team success, was a manipulation check on the instructions.

In summary, the measures, with one exception, performed quite well and can be regarded as sufficient to test the hypotheses. The inadequate measure, one of Cognitive Identification, was discarded, but a separate scale, measuring a closely related attribute, can substitute in hypotheses testing. Moreover, the bogus task appears to fulfill the requirements set for it.

Manipulation Checks

Of the three independent variables, the effectiveness of two can be evaluated by data. An analysis of variance on perceived team success, presented in Table 7, indicates that this instruction "took" very amply. The manipulation, attempting to create either Team Success or Team Failure, accounted for an estimated 21% of the variance of the perceived team success index.[1]

The success in creating the independent variable Score-No Score can also be evaluated by data. Of those 48 Ss in the Score condition, 71% responded correctly (wrote the number 6) to Question 1 of the supplementary questions,

---

[1]Many of the following analyses have been computed additionally on the components of an index separately (in this case on Question 4 and Question 5, as well as their sum) and as an analysis of covariance, with Session I SEI score the covariate. The reader can assume in all cases that the unreported computations corroborated the reported finding.

TABLE 7

ANALYSIS OF VARIANCE SUMMARY TABLE
FOR THE DEPENDENT VARIABLE
TEAM SUCCESS

Source	d.f.	M.S.	F
Competition (A)	1	.36	
Team Success-Failure (B)	1	42.31	26.74*
Score (C)	1	.64	
A x B	1	3.30	2.08
A x C	1	.38	
B x C	1	.17	
A x B x C	1	.37	
Error	87[a]	1.58	
	95		

* $P < .001$

[a]One S did not respond, hence this Analysis of Variance results from a test by general linear hypothesis rather than orthodox Analysis of Variance.

which asked S to estimate his own score. There is some indication (from post-experimental interviews) that several who responded incorrectly did indeed know their own score, but misinterpreted the supplementary question. Thus, an incorrect response here is not a sufficient reason to disregard a S's data. A Bartlett's test of homogeneity of variance was also performed. The F_{max} - value is 5.29 (*d.f.* = 47, 46; $p < .01$) suggesting that the estimates of the Score group were significantly less dispersed about their mean than were the estimates of the No Score group, the desired outcome.

While these manipulation checks are highly significant, and in the right direction, nevertheless, considering the straightforwardness of the questions asked relative to the feedback given, they are of lesser magnitude than hoped for. The perceived team success variable, for example, correlates less than .50, with the instructions given. Thus, the manipulation did not create as "clean" differences in the independent variables as intended. Perhaps the utilization of a sample whose mean intellectual capacity was closer to national averages than the present one would have yielded manipulation check data which overlapped less, and thus provided a clearer test of the hypotheses.

Randomization

Most treatises on experimental design suggest randomization as the most useful technique for equating conditions, hence its adoption in the present study. Since the random assignment to groups followed administration of the Session I measures, however, statistical tests can be made to ascertain its success. There were no differences between the "drop-outs" (the 14 Ss who were present for Session I, but absent from Session II), control, and experimental Ss on any measure except one. The Field-Dependence measure did distinguish between these groups at the $p < .025$ level ($\underline{F} = 4.45$; $\underline{d.f.} = 2, 131$), though post-hoc Scheffé tests failed to distinguish any pair-wise differences.

Directing attention to the experimental conditions, however, an analysis of variance on the Session I SEI scores, presented in Table 8 does indicate a Competition main effect, significant at the .005 level. Thus, the various conditions were not equal with regard to initial level of self-esteem. Those in the Between Sexes Competition condition had significantly higher self-esteem than those in the Within Team Competition. A difference this large will occur only once in two hundred times if accepted randomization procedures are strictly adhered to, as was done here. This failure to equate groups by randomization

has serious implications for the testability of the primary hypothesis, as will be demonstrated shortly.

TABLE 8

ANALYSIS OF VARIANCE SUMMARY TABLE
FOR INITIAL SEI SCORES.

Source	d.f.	M.S.	F
Competition (A)	1	135.38	10.13*
Team Success-Failure (B)	1	18.38	1.38
Score (C)	1	32.67	2.44
A x B	1	8.17	
A X C	1	7.04	
B x C	1	22.04	1.65
A x B x C	1	.17	
Error	88	13.36	
	95		

* $P < .005$

As the SEI was randomly halved, the question can be raised as to whether the halves were equal in terms of average score. To answer this question, the scores of the control group in each half were subjected to matched *t*-tests. The results showed an increase on the Session II version, significant at the .10 level ($t = 1.72$, $d.f. = 23$, two-tailed). The mean score on

Session I was 16.8; on Session II the mean score was 17.9. This difference could be interpreted either as inadequate equalization of halves by randomization, or as a natural increase in score when measuring self-esteem for a second time, or even as a function of the testing session for Ss (recall that Ss were singled out by E to accompany him to another room, etc.). Though the interpretation is not clear, the fact that the two halves did not yield equal scores for neutral control Ss poses problems in interpreting experimental group changes in self-esteem.

Tests of Hypotheses Using Situational Variables

Self-Esteem

The primary dependent variable is change in global self-esteem as measured by SEI. Table 9 contains the mean before, mean after, and mean gain in SEI for each group, while Table 10 presents a summary table of analysis variance on the version II (after) SEI scores. It can be seen in Table 10 that there are no effects significant at the .05 level, though the Competition by Score component does reach the .10 level. Computation of change, however, involves both a before and an after measurement. While the after measure does not contain differences between experimental conditions, the before measure does. In fact, the treatments account for an estimated 9% more variance on the before measure than they do in the after measure.

TABLE 9

MEAN SEI, VERSION I, VERSION II,
AND MEAN GAIN (II-I),
BY CONDITION.

Condition		Competition			
		Between Team		Within Team	
		Score	No Score	Score	No Score
Team Success	Version I	16.50	19.25	15.33	16.83
	Version II	17.58	16.58	15.50	18.08
	Change	+1.08	-2.67	+ .17	+1.25
Team Failure	Version I	17.25	17.92	14.75	14.50
	Version II	18.08	17.00	15.83	16.75
	Change	+.83	-.92	+1.08	+2.25
Control	Version I		16.79		
	Version II		17.96		
	Change		+1.17		

TABLE 10

ANALYSIS OF VARIANCE SUMMARY TABLE
FOR POST-SEI SCORES

Source	d.f.	M.S.	F
Competition (A)	1	14.26	
Team Success Failure	1	.01	
Score (C)	1	3.01	
A x B	1	5.51	
A x C	1	46.76	3.00*
B x C	1	4.59	
A x B x C	1	3.13	
Error	88	15.60	
	95		

* $P < .10$

Thus change is demonstrated; the change is particularly pronounced for conditions which were originally extreme in mean SEI, which now converge strongly to a common point. The difficulty in interpreting this straightforwardly is the well-known measurement bugaboo of regression toward the mean. When measurements are fallible, it is a foregone conclusion that a second observation on an extreme S will be, on the average, closer to the mean of all Ss than the first. This convergence is built into the data; it is without substantive implication.

Thus the regression toward the mean that was obtained in the present study cannot necessarily be interpreted as change in self-esteem. Had the initial SEI scores for groups been less disparate, had randomization succeeded in truly equating the groups, the regression might have been of smaller magnitude, and change in self-esteem easier to evaluate.

Here, however, this effect was so strong that a correlation of -.80 ($N = 9$, $p < .01$) was obtained between a group's (including the control) mean initial SEI and its positive change in SEI. Was regression of this remarkable degree predictable from data considering the fallibility of the measurement instruments, in which case any significant differences in change scores can be discounted as artifact, or were there some treatment

effects whose influence was to force convergence of the most extreme scores?

To answer this question, it is first noteworthy that whereas for groups, the correlation was -.80 between initial SEI and (positive) change, for individuals within groups, it was only -.30 (N = 120). A group's change score is better predicted from its initial SEI score than is that of an individual within this group. Based upon the α an estimate of measurement error, an index can be constructed which corrects the correlation of individuals' gain with initial SEI for regression--pulling out of the correlation the portion that can be predicted from knowledge of its reliability (formula in Guilford, 1936, p. 373). According to this index the correlation was -.16. The fact that the correlation diverges from zero in a negative direction implies that regression here was more than measurement error--the treatments did effect convergence. An ANOVA of raw change scores (SEI Version II minus SEI Version I score) revealed two significant components: the type of Competition main effect (F = 6.81; $d.f.$ = 1,88; $p < .05$) and the Competition-by-Score two-way interaction (F = 9.94; $d.f.$ = 1,88; $p < .01$). Since the groups differed on initial SEI, however, correction for their inequality can be obtained by utilizing it (Session I SEI) as a covariate in an analysis of covariance with raw changes as the criterion variable. This has the effect

of removing any effects on the raw change scores that could have been predicted on the basis of the pooled subjects within groups regression line and the group's mean covariate score. This analysis revealed that only the former effect, the Competition main effect, disappears when adjusted for regression, but the Competition-by-Score interaction remains (indeed, the F is slightly increased!). This latter component, then, appears to be a genuine treatment effect; it remains in an analysis of covariance where the after measure is treated as the criterion variable, and the before measure is the covariate.[2]

This two-way interaction is perhaps brought most clearly into focus by a special computation. Cronbach and Furby (in press) devote an important paper to the problem of measurement of change. According to a formula they reproduce, the reliability of the present raw difference score (SEI Version II minus SEI Version I) is a lowly .14.

---

[2]One reason that this analysis is not to be considered the primary test of the hypothesis is the fact that analysis of the covariance is predicated on the assumption of no significant differences on the covariate. This assumption is clearly violated here, though assignment of Ss to condition, it will be remembered, was completely at random. When the covariate criterion correlation is substantial, analysis of covariance is not particularly robust with respect to violation of this assumption: "In cases where differences between the covariate (cell) means . . . are relatively large, adjustment for this kind of bias is generally not based upon statistically sound principles. Such adjustments often require extrapolations which are either beyond the range of the data or in regions of the range where there are few observations." (Winer, 1962, p. 581). This covariance analysis should thus be interpreted with caution.

Raw difference scores, obviously, are not recommended as measures of change. Utilizing the principles they recommend, however, a change score formula was developed which takes into account both unreliability and test-retest regression, which, in other words, takes out of the change scores as much measurement error regression as may reasonably be permitted. (See Appendix B.1 for its rationale and derivation). The parameters of the control group were inserted where coefficients and constants were needed (though no conclusions would be altered had the parameters of the entire sample, experimental plus control groups, been substituted instead). Applying the aforementioned transformation to all the data, the ANOVA presented in Table 11 results. Again, the Competition by Score change factor is significant. Figure 1 depicts the cell means, after collapsing on the nonsignificant component, Team Success-Failure. There is a significant two-way interaction here, which does not appear to be due simply to regression but to real treatment effects.

The hypotheses predicted two out of the possible 3 two-way interactions. Since receiving a Score was expected to yield greater social comparison than not receiving one, a two-way, Score by Team Success-Failure interaction was expected of change in self-esteem. In addition, since Within Team Competition was expected to engender social comparison while Between Team Competition

TABLE 11

ANALYSIS OF VARIANCE SUMMARY TABLE
FOR GAIN SCORES AS COMPUTED BY
SPECIAL FORMULA REMOVING
REGRESSION EFFECTS.

Source	d.f.	M.S.	F
Competition (A)	1	1.27	
Team Success-Failure	1	2.30	
Score (C)	1	.43	
A x B	1	.87	
A x C	1	46.59	6.78*
B x C	1	.05	
A x B x C	1	3.26	
Error	88	6.86	
	95		

*P $<$ .05

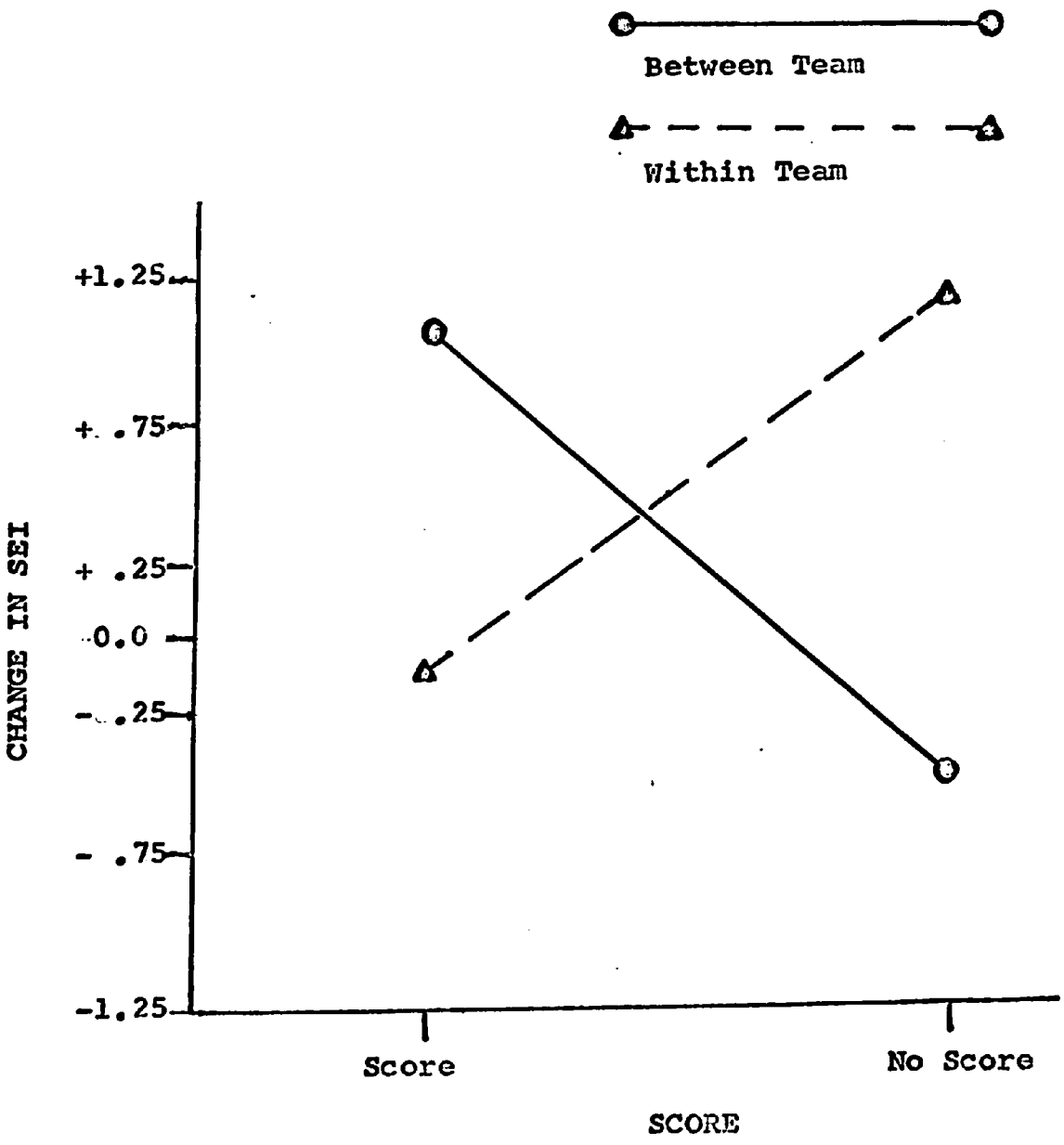


Fig. 1. Means of SEI change, computed by special formula, collapsing over non-significant component, Team Success Failure.

was expected to elicit identification, a two-way, Competition by Team Success-Failure interaction was predicted. The interaction which is significant in this analysis is, however, neither of these, predicted effects. It is instead the third possible one: Competition by Score.

In our hypotheses, Team Success-Failure was predicted to be an interactive factor, because different outcomes can be expected depending upon whether victory or defeat was experienced. In the data that was obtained, however, conditions where the Team (and, by implication for the Score condition, where the S) succeeded were little different from conditions where it failed, in terms of the impact on self-esteem. Yet though this factor washed out, the others combined to be critical. To understand how counterintuitive this result is consider for a moment only the Score conditions. Ss here learned their relative position in the class, something which should have had somewhat pronounced effects. Team Success Ss learned that they got the lowest score, by far, of the boys in the class whose tests were graded while Team Failure Ss learned that they got the best score. This variable had little effect. Instead, regardless of whether they were the best or the worst in the class, what seemed to matter is whether they were competing against the other boys, or the boys were competing against the girls. It is hard to believe that whom one is told one is competing

against has greater impact on self-esteem, than how one has performed, yet this is what the data indicates.

It is difficult to come up with a cogent post-hoc explanation which would account for this result. However, a weak one, one which turns out to have little additional support, can be fabricated.

Notice from Figure 1 that the largest part of the interaction (in fact, the only simple main effect to achieve the .05 level of significance) is the difference between the Within Team Competition and the Between Team Competition conditions where both get No Score. Suppose that Ss who get No Score disbelieve E's instruction that only half the tests were graded. Perhaps No Score Ss reason instead that they did so abominably that E wanted to save them embarrassment by not giving them feedback. For Ss in the Between Team Competition condition, having performed abominably meant, in addition to a personal failure, letting the team down. For boys of this age, letting the team down is such an unbearable event that self-esteem is damaged. To the boys in the Within Team Competition condition, however, the element of letting the team down was not present and hence self-esteem was not affected.

This "letting the team down" account, though admittedly not compelling, is one of the very few psychologically <u>plausible</u> explanations which can be reconciled with the data. It explains why the Between Team Competition

No Score condition should result in a lowering of self-esteem, but it does not account for the equally mystifying pattern of the other three points. To be sure, of all the four conditions, the one which deviates most from the mean of the other three is, by a large margin, the Between Team Competition No Score condition. However, the other three points are by no means identical, in fact they approach being significantly different themselves.

Fortunately the letting the team down interpretation makes an assumption which can be checked against additional data. The assumption is that No Score Ss believe that they have performed miserably. Since Ss were requested to guess what score they got on the bogus task after Team feedback was given, this assumption is subject to empirical verification. The mean estimate for No Score Ss was 6.88 wrong, significantly greater than that of the Score Ss (mean = 5.27, approximate $\underline{t}$ = 2.24, approximate $\underline{d.f.}$ = 65, $\underline{p} < .025$). Thus the assumption that No Score Ss would estimate a very high error score receives some support. However, the difference is not as large as would be desirable for the viability of the explanation. Also the explanation requires, since Team Success-Failure was an unimportant variable in predicting change in SEI, that it also be of little consequence here in predicting estimated score. That is, if No Score Ss believe that they performed miserably, and it is this fact coupled with their

chagrin over letting the team down which accounts for their lowered self-esteem, then it follows that there should be correspondence between how badly the Between Team Competitions conditions thought they did and how severely their self-esteem was damaged. Table 12, an ANOVA summary table (No Score Condition only) for the score S estimated he got, shows that Team Success-Failure caused significant differences in estimated score. In Table 13, which presents the sub-group means, it can be seen that in the Team Failure condition S believed he did significantly worse (8.00 vs. 5.75 wrong) than in the Team Success condition. It should follow, then according to the posited explanation, that Ss in the Team Failure No Score Between Team Competition condition should, in addition, have more greatly lowered self-esteems than Ss in the Team Success, No Score, Between Team Competition condition, since the former believe they performed more poorly. Unfortunately this implication is not supported by the data. The SEI of the former was lowered 1.20, while for the latter it fell 2.05 points, a reversal of the expected direction.

Although the letting the team down interpretation cannot be totally discarded, its performance against the data was less than convincing. Unfortunately it is the only interpretation of the data which presently seems at all plausible. Had no self-esteem changes been found, had all factors failed to yield significant results, the

TABLE 12

ANALYSIS OF VARIANCE SUMMARY TABLE
(NO-SCORE CONDITIONS ONLY)
FOR ESTIMATED SCORE.

Source	d.f.	M.S.	F
Competition (A)	1	16.33	
Team Success-Failure (B)	1	60.75	3.00*
A x B	1	.00	
Error	44	20.28	
	47		

* P (one-tailed) $<$ .05

TABLE 13

MEAN ESTIMATED ERRORS, NO-SCORE CONDITION ONLY,
BY CONDITION, AND MARGINALS

	Competition		
Conditition	Between Sex	Within Sex	Row Mean
Team Success	6.33	5.17	5.75
Team Failure	8.68	7.42	8.00
Column Mean	7.46	6.29	6.88

interpretative task would have been eased some, since it is possible to argue that performance on the task was not regarded by S as important enough for him to alter his self-esteem. The fact that though our predictions fail, our treatments did effect self-esteem change, albeit in an unexpected way, impels us to grope for an explanation. Although nothing compelling comes readily to mind, some weak and somewhat contradictory evidence can be marshalled for the preferred "letting the team down" interpretation.

Of course, the other possibility is that Type 1 error is responsible for the effect, that this is the one time in 20 that significant findings are spurious. In view of the irregularities in the data, the monumental regression effects, the significant pre-test differences, the fact that convergence was the most obvious feature of the change scores, the odds that our findings are spurious become more favorable. And since no substantive post-hoc explanations offer strong competition to this possibility, perhaps Type 1 error is the most likely explanation of the change in SEI scores.

Ability Evaluation

Though change was not computed, because an after only measurement was more appropriate, the effects of treatments on A-Eval can nevertheless be illuminating. Since A-Eval is more specific than global self-esteem, it is theoretically quite plausible that effects would

appear here, while they do not appear on SEI changes. And since we reported evidence earlier that Ss were pretty much in ignorance of their ability to "size up new situations," we can safely assume that the post-test-only measure can accurately detect "changes" in A-Eval. In Table 14 is displayed the analysis of covariance summary table for this variable[3] while the means for each condition, adjusted for variation or initial SEI, are contained in Table 15. As there were predictions regarding direction, one-tailed tests were performed (since all are 1 d.f. tests, the square root of $F = t$, and t-tables are directional). It can be seen that the Score by Team Success-Failure, two-way interaction is highly significant, a predicted finding.

Thus, higher self-evaluations predominated in the Score, Team Failure and No Score Team Success conditions while lower self-evaluations were found in Score Team Success and No Score Team Failure conditions. We define social comparison effects as the occurrences of raised self-appraisal following M's failure, and lowered self-appraisal following M's successes. Thus it appears that social comparison effects were more dramatic in the Score condition, where S had complete feedback, than in the No Score condition, where S was less certain of his own performance, as we predicted in hypothesis 1.

---

[3]Initial level of self-esteem is the covariate. This correlates .15 with A-Eval, high enough for there to be gains over ANOVA, low enough that the comments of footnote 2, this chapter, are of little force.

TABLE 14

ANALYSIS OF COVARIANCE SUMMARY TABLE FOR A-EVAL, UTILIZING INITIAL SEI AS COVARIATE

Source	d.f.	M.S.	F
Competition (A)	1	.51	
Team Success-Failure (B)	1	.41	
Score (C)	1	.02	
A x B	1	1.99	1.23
A x C	1	.18	
B x C	1	46.67	28.95*
A x B x C	1	1.79	
Error	87	1.61	
	94		

* $P < .001$

TABLE 15

CELL MEANS FOR A-EVAL, ADJUSTED FOR DIFFERENCES ON THE COVARIATE INITIAL SEI, BY CONDITION

Condition	Competition			
	Between Team		Within Team	
	Score	No Score	Score	No Score
Team Success	6.75	8.00	6.12	7.74
Team Failure	7.73	6.71	8.22	6.48

Note: The higher the cell entry, the more positive the evaluation.

Though the other predicted interaction failed to attain conventional levels of significance, the pattern of the means in Table 15 prompted a test of the simple interactions.

For the Score conditions alone, the predicted Competition by Team Success-Failure effect attained marginal significance. ($F = 2.36$; $t = 1.53$; $d.f. = 87$, $.10 < p < .05$). For the No-Score conditions, however, the effect (Competition by Team Success-Failure interaction) failed to materialize. In general, speaking of the Score means only, Team Failure yields higher A-Eval than does Team Success. But the difference is substantially larger in the Within Team Competition (6.12 vs 8.22) than in the Between Team Competition condition (7.33 vs 6.75). According to our definitions, this implies that Within Team Competition, where S presumably sees M as a rival, promotes more social comparison effects than does Between Competition, where S presumably regards M as an ally. This is in accordance with Hypothesis 2. However, directing attention to the No Score means in Table 15, the same tendency for Within Team means to be more extreme than the Between Team A-Evals does not prevail.

Considering the A-Eval results, it appears that the hypotheses are well, if not perfectly, supported. Score does interact in the predicted way with Team Success-Failure; Competition does interact in the expected way with Team Success-Failure, but only when Ss received a Score for themselves.

There are several reasons, apart from the partial failure to substantiate one of the hypotheses, why the above results should be regarded with some caution.

The A-Eval measure figuratively asks "How do you feel about your ability?" rather than, "How do you feel about yourself?" as does SEI. To demonstrate unequivocally vicarious sharing in triumph, the "belong to good team" interpretation of the identification hypothesis, we require that the second but not the first of these questions be answered in a positive way. We are interested not in whether or not S feels he could perform as M did, but rather in whether he shares in the affective consequences of the performance, e.g., joy in victory, dispair in defeat. Yet positive results with A-Eval do not have this meaning; what S evaluates is how he could or did perform. Thus, in one respect, the theoretical predictions are not well tested by A-Eval.

Secondly, A-Eval is a good deal more specific than is change in self-esteem. We can distinguish three levels of generality for which data is available. The first is the perception of the actual score obtained, the second is A-Eval, and the third is change in SEI. It is obvious that score obtained is very close to the manipulation (in fact served as a manipulation check for Score Ss), but A-Eval is but slightly more remote. Consider, for a moment, just the Score conditions. Ss in the Team Success condition learned that they had the lowest score in the

class, while those in the Team Failure learned that they got the best score in the class. The fact that the former would have lower A-Eval than the latter is sufficiently apparent that had this not been the case we'd question whether the instructions had been properly understood. In other words, for these Score conditions, the A-Eval is theoretically quite close to the manipulation, so that the confirmatory results have something of the nature of a manipulation check. For the No Score conditions, things are much less obvious. When not given a score, an S whose classmates did well thought he had more ability than an S whose classmates did poorly. This can not be a case purely of understanding the feedback but instead must be a case of introjecting on the basis of similarity to M. This finding replicates that found by numerous writers (Burnstein, et. al., 1960; Shupp, reported in Diggory, 1966; Cutick, 1962). For the No-Score conditions, the results with A-Eval are, then, of substantial interest.

What of the other two-way interaction, Competition by Team Success-Failure, which obtained only for Score Ss? This was predicted by Hypothesis 2, and with the exception of Ferguson (1958), not investigated by any study. Moreover, the prediction is of a more non-obvious and therefore, useful, character. However, and this is the final reason for caution in considering the A-Eval results as confirmatory of the hypothesis, the above interaction can be explained in an alternative way. Recall that we had

predicted that when Between Team Competition prevailed, the S was thought to regard M as an ally, but M was presumably regarded as a rival in Within Team Competition. This ally vs rival notion led to the hypothesis that identification effects were more likely in the Between Team condition.

The alternative explanation of the simple interaction, Competition by Team Success Failure depends upon the fact that the instructions for the Between Team Competition condition introduced the notion of another Team, besides the Team of "Boys in the Class," namely that of "Girls in the class". The instructions for the Within Team Competition condition, by contrast, made no mention of any other than the Boys in the Class. For all the Team-Failure condition, the Boys had low scores, but for the Between Team condition they indeed were "so low that I"d be surprised if the girls in the class are not able to do better than this." Those Ss receiving a Score learned that they did the best of the boys, but this was not so sterling a performance for the <u>class as a whole, since the girls probably do better than this</u>. Their position as the superior boy is diluted by the real possibility that they are also inferior to many of the girls. The <u>Within</u> Team Competition, Score, Team Failure Ss, have only the boys to compare scores with. Their relative superiority is <u>not</u> diluted by a comparison of boys with girls. It follows then that the A-Eval of Within Team

Competition, Score, Team Failure Ss will be higher than the A-Eval of Between Team Competition, Score, Team Failure boys. By an analogous but opposite argument, Between Team Competition, Score, Team Success Ss should have poorer A-Evals than Within Team Competition, Score, Team-Failure Ss.

To say this another way, considering only Ss in Score conditions, those in Between Team Competition have their superiority or inferiority to boys tempered by the comparison of boys to girls, while those in this Within Team Competition condition do not. This implies that the A-Eval of the former should be less extreme than that of the latter, a condition met by the data presented in Table 15, and one which would result in the Competition by Team Success-Failure two-way interaction that was predicted and obtained for the Score data.

Interestingly, the above argument, which we shall call the "absolute superiority" interpretation would not apply to Ss in the No-Score conditions as readily, since they do not know their relative standing within the group of boys. Hence, according to this interpretation, the interaction would appear only in the Score, not in the No-Score conditions. In the ally vs rival argument, however, which formed the theoretical basis for Hypothesis 2, the prediction was that the interaction would not be confined to the Score conditions, but would apply equally to Score and No Score. Since the alternative "absolute superiority"

interpretation actually fits the obtained data with more accuracy than our hypothesized "ally vs rival" argument, the "absolute superiority" interpretation can not be disregarded. We must reserve as a possibility the fact that the dynamics we posited in our theoretical development are not responsible for this effect.

Importance

One of the variables used as a measure of ego-protection, theoretically a variant of self-appraisal, was the importance S attached to the ability, "being able to size up new situations." In the instructions, Ss were told that E regarded this ability as extremely important, but the Supplemental Questionnaire gave S a chance to rate its importance himself. By and large, they agreed with E (the mean rating was 4.12 which corresponded most closely to the phrase "pretty important") which was encouraging. Had it been too low, no reappraisal of self could have been anticipated, for what does it matter if I have little (or much) of an ability I consider trifling.

Having been thus encouraged, attention was then directed at the manner in which importance varied systematically with the experimental conditions. This variable is largely orthogonal to the other dependent variables reviewed above, so that results concerning it may well add to our understanding of the processes at work in the experimental situation.

It was expected that the importance S attached to doing well on the task would operate much as a perceptual defense mechanism. That is, when self-esteem was threatened because of performance, the threat could be reduced by perceiving the task as unimportant. Self-esteem can escape unscathed if the challenge is dismissed as rather trivial. Where self-esteem is being enhanced, however, greater gains can be made the greater the perceived importance. Thus, self-esteem and perceived importance were hypothesized to react in similar fashion, and the predictions using Importance were identical to the predictions for the previous measures of self-appraisal. Table 16 presents the analyses of variance for the importance S attached to the ability. Note that the Competition by Team Success-Failure interaction is significant. The raw means are depicted in Figure 2. Where there was Between-Team Competition and the team failed, and where there was Within-Team Competition and it succeeded, Ss interpreted the ability in question as relatively more important than in the other conditions.

With regard to the Competition variable, it was predicted that for Between Team Competition, where the success of the Team was desirable, the Team's failure should cause S defensively to perceive the task as unimportant. For Within Team Competition, on the other hand, the operation of social comparison processes should serve to render desirable the failure of the team. Where it

TABLE 16

ANALYSIS OF VARIANCE SUMMARY TABLE
FOR S'S PERCEIVED IMPORTANCE
OF TASK

Source	d.f.	M.S.	F
Competition (A)	1	.01	
Team Success Failure	1	2.34	2.11
Score (C)	1	.84	
A x B	1	7.59	6.85*
A x C	1	1.76	
B x C	1	1.26	
A x B x C	1	.26	
Error	88	1.11	
	95		

*$p < .02$, two-tailed

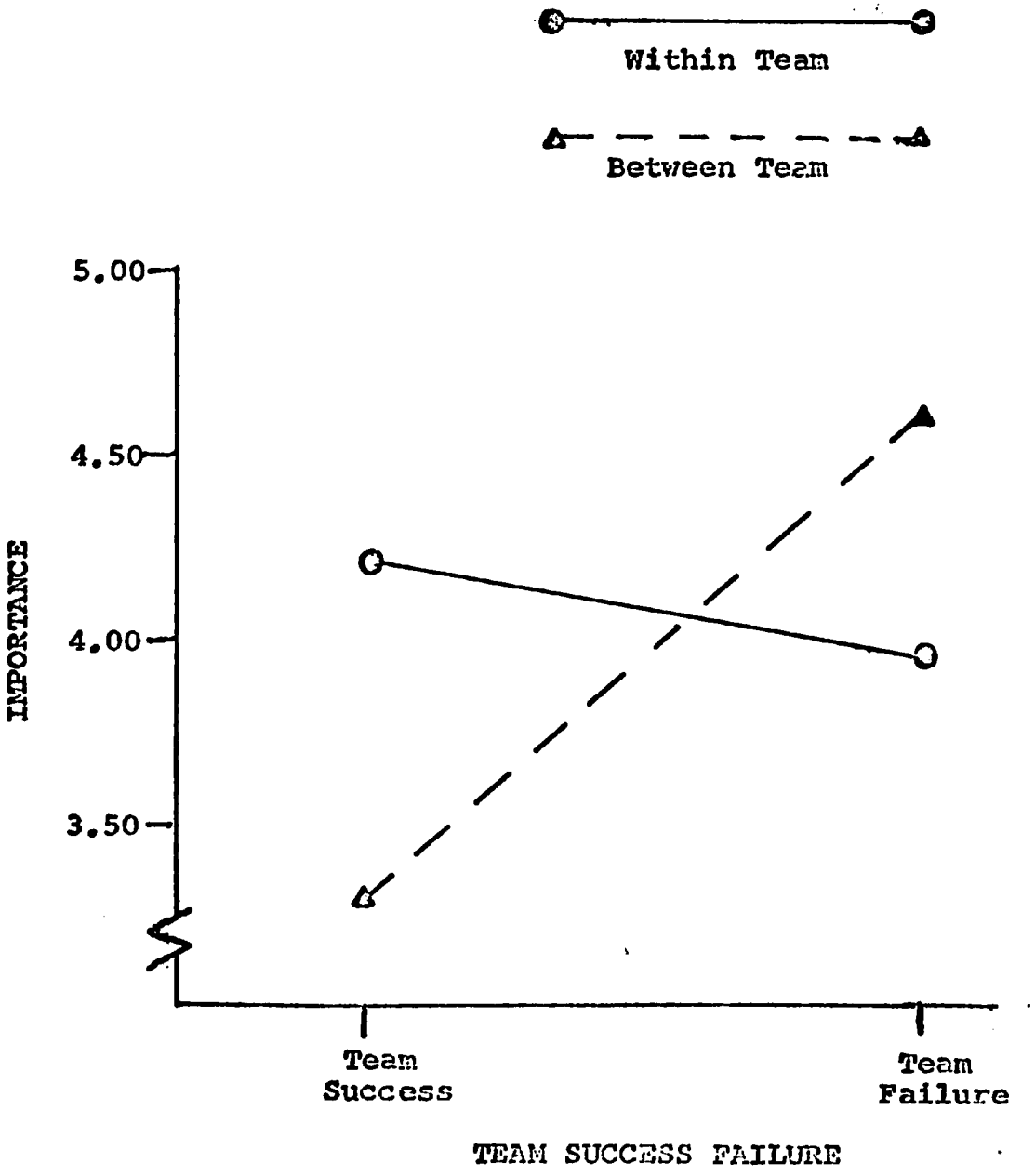


Fig. 2. Means for the perceived Importance of doing well, after collapsing on the non-significant component, Score.

succeeds, however, the perception of the task as unimportant could function to protect self-esteem. These predictions amount to a two-way Competition by Team Success-Failure interaction, as found. However, the direction is reversed! Those conditions which were predicted to be highest on perceived importance were instead significantly lowest. Moreover, the Score by Team Success-Failure two-way interaction, which also was predicted (and strongly appeared when the dependent variable was A-Eval) failed to achieve significance for perceived importance.

One possible explanation which was entertained to account for this reversal was the "impact of the unexpected" interpretation. Suppose that Ss with high self-esteem expect themselves (or that with which they are associated, i.e., their Team) to prosper, while inferior feeling Ss expect to fail. Feedback which confirms these expectancies has little impact (is treated as unimportant) but that which violates the expectancy is regarded as important. Note that, because of the failure of randomization to equate groups, the Competition manipulation was significantly confounded with initial level of self-esteem. The Between Team Competition condition was high in self-esteem, and hence, may have expected Team Success. The Within Team Competition condition, on the other hand, was low in self-esteem and perhaps expected Team Failure. For the former, Team Success would have confirmed the expectancy, and hence the feedback would have had little impact

or importance as compared with Team Failure. For the Within Team Competition condition, the opposite held. By translating the Importance dimension into an Impact dimension, by assuming that Team Success or Failure is expected on the basis of prior self-esteem, and by demonstrating that prior self-esteem, because of randomization failure, varies with Competition condition, the reversal can be accounted for.

To test this interpretation, analysis of covariance was performed on the Importance response measure, using prior self-esteem as the covariate. This procedure, in essence, equates the groups with respect to prior self-esteem, removing the confounding. If the above "impact of the unexpected" argument is valid, this procedure would have eliminated the interactions which the analysis of variance revealed.

The results of the analysis of covariance were little changed from that of the analysis of variance. The F-values for all the components of variance remained almost constant, and Figure 3, containing the means, adjusted for variance on Version I SEI, indicates that the interaction was not removed, but, rather, became more balanced among the Competition conditions with regard to the effect the Team's Success or Failure had.

Assuming, then, that this interaction is genuine and not Type I error or artifact, how can it be interpreted? A large accumulation of literature, for example,

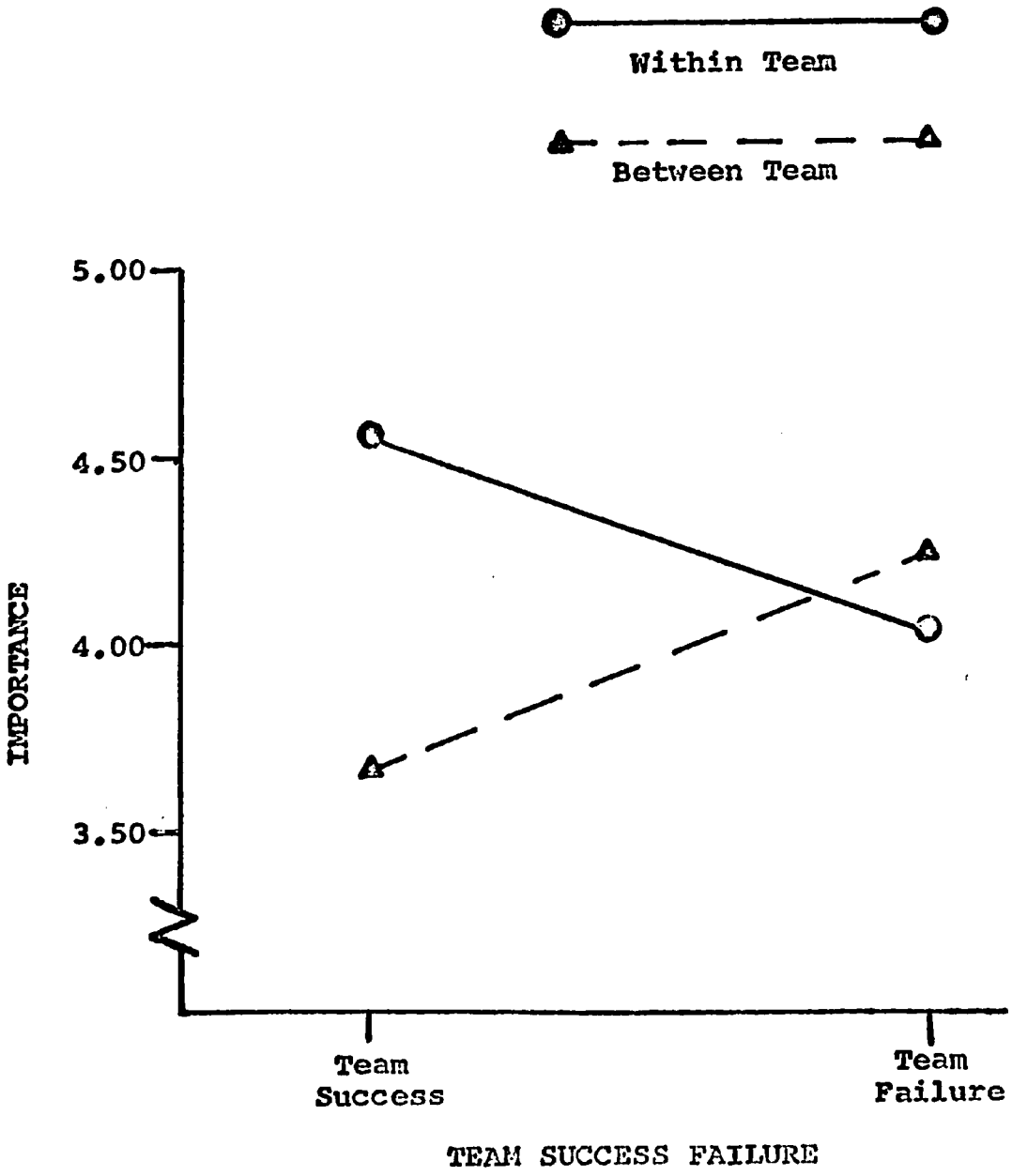


Fig. 3. Means, adjusted for variation on Version I SEI, for the perceived Importance of doing well, after collapsing on the non-significant component, score.

in dissonance, manifest anxiety, achievement motivation, etc., gives evidence that threat to self-esteem, such as engendered by eminent failure, evokes ego-defensive mechanisms. Amongst these mechanisms is that of denial; denying the importance of the threatening event can serve to prevent the painful task of lowering self-esteem. It was in accord with this reasoning and literature that the present hypothesis was formulated.

Yet, often the ego-defensive interpretation in past research failed to be substantiated. It is a rather elusive effect to obtain. The personality variable of repressor -sensitizer was developed by Byrne (1961, 1964) to account for this difficulty. For repressors, ego-defense mechanisms are rapidly called into play when self-esteem is threatened. For sensitizers, however, ego-defenses are not called into play, but instead, the threatening information is regarded with heightened attention. Feedback of failure, for example, might be perceived by sensitizers as greatly important. Were our sample to be composed predominantly of sensitizers (for which there is neither confirming nor disfirming evidence) the obtained result would be more plausible. For Between Team Competition, where Team Failure would be more threatening than Team Success, sensitizers would be predicted to react with heightened awareness, i.e., award more Importance, to the former. For Within Team Competition, where the others are presumably rivals, the greater their success,

the more S's ego is threatened. This results, for sensitizers, in increased attention, and the attribution of increased importance to the information. Such was the data.

One difficulty in reconciling this argument with the data is that it requires that the greater the ego-threat, the greater the importance S attaches to the information. In the present experiment, surely the greatest ego-threat that occurred was in the Score Team Success (regardless of Competition) condition. Here S learned he performed the worst in the class. As compared either to the Score Team Failure (where S learned he did best in the class) or the No Score conditions (where he could not be certain of his own performance) threat was at a severe level. We would expect, if the sensitizer effect were a valid explanation of the obtained interaction, that it would also involve, in some way (either as a main effect, interaction, or simple interaction), the Score variable. The fact that the obtained significant component of variance, Competition by Team Success-Failure, is one that by this explanation would be less likely than another which did not attain significance dampens the credibility we can attach to the sensitizer argument.

Liking for Classmates

The other variable of ego-defensiveness utilized was the attraction S expressed for his male classmates

after full feedback. Ss were asked to evaluate both their classmates who were their close friends, and those who were not.

It was speculated that the affect experienced by Ss as a result of success and failure would radiate to expressions of magnanimity and affection, or frustration and externally directed punitiveness, respectively, for fellow man. As Isen (1970) put it, there is a warm glow accompanying success. In the case of the present study, this is of particular interest since success and failure are actually a direct result of the performance of classmates. Thus it was expected that the pattern of results might be complex, but it was reasonable to hypothesize some effects of success and failure on liking for classmates.

No significant effects were obtained, however. As the two questions, Supplementary questions 7 and 8, did not correlate positively, their combination into a scale[4] was without justification. Thus, taking the questions separately, no treatment effects whatever were found.

As this variable was postulated to be related in intricate fashion, if at all, to self-reappraisal, the

---

[4] In fact, a scale based on the sum of their Z-scores was planned, since as anticipated, the variance was smaller in the first, (F_{max} = 2.08; d.f. = 95,95; $p < .01$) and hence deviations count more.

failure to find any significant treatment effects on it can be regarded with little concern. It was included much as a long-shot or after-thought, and is probably of little implication with regard to the validity of the hypotheses.

Summary: Tests of Hypotheses Using Situational Measures

In sum, the three situational variables, the independent variables in the design, were looked at with regard to their effect on the following dependent variables: changes in self-esteem, A-Eval, perceived importance of the ability, and liking for classmates. Significant treatment effects were observed for each of the first three variables, but only for A-Eval was it substantially as predicted. For both Importance and changes in self-esteem, post-hoc interpretations were offered which could account, in part, for the data. In both instances, other implications of the arguments were not at all, or not completely, substantiated by the data. Because of irregularities in the measurement, it was proposed that the self-esteem change finding might well be attributable to Type I error.

The A-Eval result was largely as predicted by the hypotheses. Yet it was advised that the findings be regarded with some caution, in large part because an alternative interpretation accounted for the data at least as well as the hypotheses.

Tests of Hypotheses Concerning Individual Difference Variables

Most tests of the individual differences hypotheses concerned the degree to which the individual identified or socially compared himself with the other members of his "Team," as a function of his personal characteristics. The operationalization, i.e., the measure, of identification effects and social comparison effects was intended to be the direction and degree of change of self-esteem. For example, Hypothesis 4, regarding the variable of initial level of self-esteem, posited that greater identification effects would be found for those low, than for those high in self-esteem. Identification effects were to be measured by self-appraisal changes as a function of the Success and Failure of M.

In the previous section, the data on SEI changes made it clear that, whatever accounted for the treatment effects, in all probability it was neither identification effects nor social comparison effects. Thus, it would be inappropriate to carry out the planned tests of the individual difference variables correlating them against SEI changes.

The variable of A-Eval, however, reacted largely as hypothesized to the treatments. It remained reasonable, therefore, to regard A-Eval, as a measure of self-appraisal, also as a measure of the prominence of identification

effects and of social comparison effects. Inasmuch as the effects with A-Eval, were, however, not completely as expected and could be explained in alternate ways, its usage as an index of identification effects should be treated simply as suggestive, rather than as unequivocal evidence for or against the hypotheses.

Thus we shall correlate A-Eval against the individual difference variables to obtain tests of the hypotheses which concern them. One drawback in this procedure was the specificity of A-Eval as a measure of self-appraisal. This was especially apparent in the Score condition; when S gets feedback about his own performance, we can assume it is primarily this feedback, rather than anything related to identification effects, that is reflected in his self-evaluation. It was appropriate in most instances to exclude the Score conditions from tests in this section.

Vicarious processes in general, not solely identification effects, can be predicted to vary as a function of the personal characteristics studied. It was possible to employ, in some instances, the variable of the number No Score Ss estimated they got wrong as a measure of the degree to which S introjected, or took on for himself, M's characteristics. (For Score Ss, this variable simply indicated whether they heard, understood, and believed E's feedback that they got 6 wrong.) Since Ms received Scores, the number S estimated he got wrong indicated the

degree to which he felt his performance resembled Ms'. Specific inferences using this variable, which reacted in reasonable fashion to the situation variables, are explained where they occur.

It is worth noting that the correlation of after-SEI with A-Eval is not high, though positive, .25. Much of this correlation, however, is due to the effect of global self-esteem influencing the evaluation of ability, rather than vice-versa. The correlation of before SEI (a chronologically prior, and hence antecedent variable) with A-Eval is .15. A partial correlation of A-Eval with after-SEI, controlling for initial SEI, is .20. Thus, change in self-esteem is not at all strongly influenced by the favorableness of one's evaluation of one's ability.

Hypothesis 3: Strength of Identification

Due to our inability to utilize SEI changes as a measure, it was thought better to modify the prediction somewhat, in order to test it as well as possible. The Competition variable was collapsed.

The hypothesized effect of strength of identification is represented by the signs in parentheses in Table 17. In those cells where identification effects seem to have been found (the No-Score conditions), strength of identification is expected to increase the effect, i.e., correlate positively with A-Eval where the Team Succeeds, and correlate negatively where it fails. In the Score

condition, on the other hand, the attraction-to-team aspect of strength of identification should lead to the findings of Festinger, Torrey, and Willerman (1954), namely negative correlations with A-Eval where S performed relatively poorly, positive correlations where he performed well. Recall that the interpretation they made was that those Ss who were highly attracted to their group, seemed to be more greatly concerned with their relative position, i.e., they engaged more in social comparison processes. The correlation coefficients are also given in Table 17. None are significant, and only half (the smaller r's) are even in the right direction.

TABLE 17

CORRELATION COEFFICIENTS FOR STRENGTH OF IDENTIFICATION WITH A-EVAL (HYPOTHESIZED SIGN IN PARENTHESIS) BY CONDITION

Condition	Score	No-Score
Team Success	-.07 (-)	+.24 (+)
Team Failure	-.17 (+)	+.35 (-)

The Score finding can be considered to directly contradict that of Festinger, et. al., particularly given the similarity of the procedures in the Score condition to those they employed. This failure to replicate might

be due to the different characteristics of the samples employed. While Festinger et. al., used college students, a rather low intelligent six grade boy was the S in the present study. Perhaps their result is restricted to Ss of college student mentality.

It is of interest that though the predicted "interaction" did not occur, a "main effect" could be found with regard to the correlation. Thus, the correlation of strength of identification with A-Eval for the No Score group was significantly positive (+.43, $p < .01$) and significantly greater than that of the Score group ($Z = 2.28$, $p < .03$, both tests two-tailed). Again, as was the case with change in self-esteem, the Team Success-Failure variable--which was expected to be pivotal--washed out, while another variable, in this case, Score, was crucial.

In retrospect, it doesn't appear surprising that for Score Ss the correlation of A-Eval with strength of identification hovered near zero. For Score Ss, as we earlier argued, it is the score they receive, and its relation to the Ms' scores that shapes A-Eval.

The positive correlation obtained for the No Score Ss, regardless of Team Feedback, is more puzzling.

One possible explanation which was entertained to account for the finding involved a possible artifactual relation. No Score Ss, having no feedback information to the contrary, perhaps assume that they have about as much ability to size up new situations as they have other

abilities. Their global evaluation of their other abilities is what is meant by the term self-esteem, so what we are proposing is that there is a high correlation between self-esteem and A-Eval for No Score Ss. Suppose, furthermore, that strength of identification is also highly related to self-esteem, since we assume that boys who evaluate "Boys" highly, also evaluate themselves this way. (Recall that the measure of strength of identification we are relying on is "affective identification", which in essence, was a rating of the concept "Boys" on the evaluative dimension.)

Given the above two assumptions, the conclusion follows that the correlation between A-Eval and strength of identification must be positive, because of the positive relationship of the two to self-esteem. A partial correlation controlling for initial SEI reduced the correlation somewhat, to .32 (from .43). This partial correlation however, was not small enough to give full credence to the putative artifactual relation to self-esteem. In fact, the correlation of SEI to A-Eval for No Score Ss is slightly below (.39) the correlation of A-Eval to strength of identification, the correlation for which we are attempting to account.

Nevertheless, the artifact of self-esteem argument remains so compelling that it is difficult to relinquish. And it is difficult to interpret the result in alternative ways with any real confidence.

Hypothesis 4 and 6: Initial Level of Self-Esteem

The test of the hypothesis concerning effects of initial level of self-esteem on identification effects or social comparison tendencies has been partially compromised by the failure of randomization and the failure of the primary dependent variable. Within a condition, the range of SEI is restricted - some conditions contain Ss with generally high self-esteem, others with generally low. Thus the within condition correlation which results is biased - probably toward minimizing the strength of the relationship. However, if we collapse on the Competition variable - one whose effect is marginal in any case - this difficulty is avoided.

Doing so, it remains reasonable to ask whether high SEI's are more likely to manifest identification effects, i.e., to attribute as much ability to themselves as their teammates displayed, than lows are. The primary test of this hypothesis will be given by correlating A-Eval to initial levels of self-esteem (Session I SEI score). We predicted that the correlation would be negative in the Team Success condition and positive in the Team Failure condition. (The low SEI Ss were predicted to manifest the greatest identification effects.)

The correlations were +.17 for the Team Success and +.16 for the Team Failure conditions, almost identical.

The hypothesis is clearly not supported when A-Eval is our measure of self-reappraisal.

In addition, however, a second approach to answering this question was attempted; it concerned the estimated task scores of No-Score Ss after team feedback is given. In the Team Success condition, introjecting Ss are those who give the lowest estimations of numbers of incorrect answers; in the Team Failure conditions, they are those with the highest estimates of errors. Do these introjecting Ss tend to be high or low in initial self-esteem? In part, the answer is obscured by the fact that there is a bias for high SEI's in general to predict lower error scores than low SEI's. However since Ss also estimated their errors before team feedback was given (in Session I) and since we can assume that the effect of general level of self-esteem on estimates of errors operates equally for both estimates, we can correct for this bias by partialing out the error estimates before team feedback was given, i.e., the Session I estimates. The resulting partial correlations are -.10 in the Team Success conditions, +.24 in the Team Failure Condition ($N = 24$; $p < .15$, one-tailed difference tests) which are consistent in showing that high SEI Ss are those most obviously introjecting. This is evidence opposed to the hypothesis but weakly substantiating Stotland and Cottrell's findings (1961, 1962). The relationship between tendency to introject and self-esteem is obviously a complex one, since several studies,

including the present, have found the counter-theoretical result that high self-esteems are more likely to introject, while other studies have obtained contrary results. This question awaits careful research.

The data to be reported next concerns the self-fulfilling prophesy character of self-esteem. Does a person high in self-esteem think more highly of his performance - even after full feedback - than a person whose self-esteem is low?

The answer has a number of facets. Initial SEI does correlate significantly (r = .21; N = 134; $p < .05$) with guessed score before feedback, but not with guessed score after feedback (r = .05; N = 96). With A-Eval, across all conditions, the correlation of initial SEI is .15 (N = 96, N.S.). However, with one of the two questions which comprised the A-Eval index ("How good are you at sizing up new situations?"), the correlation was .24 (N = 96; $p < .05$). The tendency held up in the Score conditions as well as in the No-Score conditions. There is weak supportive evidence here; if two people are given identical performance feedback, the one with the highest self-esteem has a slight tendency to interpret it the most favorably.

When SEI is correlated against change in SEI, the correlation, after correction for measurement error regression, is -.16, as stated earlier. This indicates that despite the fact that high SEI's tend to interpret

their feedback favorably, they tend to fall in SEI; low SEI's,who interpret their feedback negatively, however, tend to gain. This seeming contradiction, however, is resolved when considering that the treatment was not constant for all Ss, and because of randomization error, high SEI Ss tended to be assigned to different treatment conditions than low SEI Ss, a confounding of initial SEI with treatment. Moreover, the particular treatments that highs received seemed to be treatments which lowered SEI, while the lows generally got treatments which raised SEI. Finally, the change in SEI is only marginally related to A-Eval in any case.

Hypothesis 5: Field-Dependence

The hypothesis was that field-dependents (i.e., those with low scores on the instrument) were more likely to engage in identification processes. Thus, the prediction was made that a negative correlation in the Team Success, a positive correlation in the Team Failure condition would be found between Field Dependence and measures of self-appraisal.

Again the planned tests, correlations against changes in SEI, were rendered inappropriate by the failure of the latter, the primary dependent variable, to yield hypothesized results in the testing of situational variables. As a compromise, however, computations were carried out with A-Eval, and with the number of problems

S estimated he got wrong subsequent to team feedback in the No-Score conditions. (In the Score Conditions, these variables were so theoretically close to the independent variables or manipulations, that they functioned largely as manipulation checks. In the No-Score conditions, however, in which introjection-identification processes are presumed to operate, individual differences therein can be more legitmately studied using the above variables.) The correlations of Field-Dependence and number guessed wrong, after partialing out the guess in Session I, was -.05 and -.13, for the Team Success and Team Failure conditions, respectively. These correlations are so weak and so close to one another, that the variable of Field Dependence cannot be said to relate to introjection as measured by guessed scores. The correlations of A-Eval and Field-Dependence, however, were .12 and .40 (the last only is significant at .05, $N = 24$), for the Team Success and Team Failure condition, respectively. (The reversal of sign simply indicates that the indices have reversed scoring directions, i.e., number of errors is inversely related to A-Eval. For all coefficients, however, the relation of Field-Dependence to perceived ability is positive.) There seems, then, to be a tendency for Field Independents to feel they did well. As Field-Dependence correlated with self-esteem, however ($r = .22$, $N = 134$, $p < .05$), which also correlated with A-Eval, partialing it out lowered the correlations to .03 and .39 respectively, the last remaining

significant. The difference between the two approaches significance ($Z = 1.24$; $p < .11$). The partial correlations tended to render the hypothesis more palatable. Although the Team Success correlation was not quite negative, as predicted, it was clearly lower than that for Team-Failure, as predicted. To the degree that we can accept A-Eval as a measure of tendency to introject (and neglect number guessed wrong, where the data were far less dramatic, though in the right direction) the variable of Field-Dependence can be said to show real promise as a correlate of the tendency to identify. This is encouraging, because a number of other such variables, both in this and previous studies, have failed to relate systematically to identification.

CHAPTER V

CONCLUSIONS

Summary

In part, the burden of this paper was to point out a conflict between two theories, each rich in supporting literature. The hypothesis of identification effects states that, in general, people take on characteristics and reactions of others with whom they identify. The hypothesis of social comparison processes states that people compare their attributes to that of others whom they use as referents to evaluate their own. With regard to self-appraisal, the hypotheses make more particular statements. Identification processes imply that the more the (identificatory) other succeeds, the greater S's self-appraisal; the worse the other's failures, the more severe the decrease in S's self-appraisal. With regard to the (referent) other's performance, social comparison processes imply the direct opposite relation. There are many instances where an other is at once a referent and an identificatory other. (Indeed, most identificatory others serve additionally as referents, though the reverse is not true.) When this is so, the predictions of the two processes come

into inescapable conflict. When other succeeds, identification effects dictate an increase, social comparison effects a decrease, in self-appraisal. When the other fails, the opposite is predicted.

A number of variables were posited, which according to the predictions made, should influence the resolution of the conflict in favor of one or the other of the counterpoised theories. An experiment which manipulated or measured the proposed variables was performed, and the relation between the variables and the two possible effects were observed. The dependent measures of the effects were measures of self-appraisal, specifically: change in self-esteem, self-evaluation of the tested ability, the importance attached to the ability, and the expressed liking for classmates.

Situational Variables

With regard to the major dependent variable, change in self-esteem, the fact that randomization failed to equate the conditions with regard to initial level of self-esteem rendered the variable nearly useless. Corrective statistical controls were applied, however, in an effort to salvage the instrument to the degree possible. When this was done there was not only no expected effects caused by the manipulations, but also no evidence either of social comparison nor of identification effects. The operational definitions of the two effects pivoted upon

whether the Team Succeeded or Failed, yet the data did not "break" differently depending upon the Team's performance. Some of the experimental situations resembled those studied in previous research with identical or similar measures. For example, the Score Within Team Competition conditions can be considered an attempted replication of the Morse and Gergen (1962) study, which, using SEI, obtained social comparison effects. The No Score Between Team Competition conditions resembled the paradigm of Burnstein, Stotland, and Zander (1961) which was able to obtain identification effects. (Identification effects in the latter were confined to the conditions where M succeeded; here S gained in self-appraisal. In the present study, this is the condition with the most pronounced reduction in self-appraisal.) In the Morse and Gergen study, the Ss came from markedly different populations, but in the case of Burnstein, et. al., the Ss were, as in the present study, grade-school boys. Thus, effects on the variable of self-esteem, which had been found in previous studies, failed to be replicated under the similar conditions, here.

It was not the case that no significant differences attributable to treatments were found, however. An unpredicted, difficult to account for, interaction effect was obtained. The proferred interpretation was not totally congruent with the available data, however. The possibility of Type I error, randomly caused rather than treatment caused differences, was deemed a reasonable alternative

explanation, because of various irregularities in the data.

With regard to the secondary dependent variable of A-Eval, a pattern of results nearly identical to that hypothesized was obtained. Though support for the situational variable hypotheses was not obtained with change in self-esteem as the response measure, they were substantiated when the measure was A-Eval. This essentially replicated the results found by Ferguson (1958). It appears, then, that identification and social comparison effects are obtainable only for specific self-evaluations. There is, however, no radiation or spread of effect from evaluation of specific ability to evaluation of total self.

It was advised that these confirmatory results be considered with caution, since in some aspects they were quite close to the manipulations, and, hence, exceedingly likely. In other respects, the results were also explainable by alternative accounts, which had little to do with the dynamics of identification and social comparison.

A third dependent variable, perceived importance of the ability, also was analyzed. A predicted interaction was obtained, but examination proved it to be in the opposite direction from expected. It was speculated that the results found could be attributed to the possibility that the sample was composed predominantly of

sensitizers. If this were the case, information which was the more ego-threatening should be regarded as the more important, the obtained finding. A major implication of this account, that the Score variable should have additionally made a difference, was not substantiated, however.

The final dependent variable was liking for the boys in the class, a variable included somewhat as an afterthought. There was found to be no differences whatever between treatment conditions on this variable.

To recapitulate the hypotheses concerning situational variables, and their status in light of the experimental data:

(1) The more surely S knows of the acceptability of his own performance, the more likely he is to engage in social comparison, rather than identification, processes. This was tested by the significance of a two-way, Score by Team Success Failure, interaction. Of the four dependent variables, the predicted effect occurred only for A-Eval. It appears that, as in the bulk of past research, evaluation of one's specific ability tested is the one variable most likely to be affected by one's own, or by an other's, performance. Perhaps the other dependent variables are too indirectly related to this to be influenced by the kind of phenomenon we are able to

simulate in a laboratory study. We were able to obtain the basics, but we weren't able to obtain the frills. Perhaps A-Eval has a more cognitive, information-processing, character (where S must only update his information on the basis of the new feedback) while self-esteem has a more emotional and affective character. To evaluate an ability is somewhat impersonal; self-esteem means to evaluate <u>oneself</u>. Ego-involvement, and resistance to change, is apparently far greater in the latter case.

(2) How the other is regarded, as an ally or as a rival, has implications for the resolution of the conflict between the processes. We are likely to socially compare in those situations where we regard the other as our rival but we should identify when we see him as an ally. Operationally this implied a two-way Competition by Team Success Failure interaction. This prediction, at least in part, was confirmed for two of the four dependent measures: perceived importance and A-Eval. For the former, however, the means were reversed from expected; for the latter the interaction was only obtained in the Score, not in the No Score condition.

We must regard this hypothesis carefully. We cannot discard it, for the post-hoc interpretation made to account for the reversal for perceived importance <u>still</u> affirmed the threatening quality of team or personal failure, and the uplifting quality of success. Yet we cannot enbrace it, for to do so on the basis of a post-hoc

interpretation would be premature. We cannot, similarly, regard the simple interaction confirmation (which was only marginally significant in any case) as solid support, for it could be interpreted as readily elseways. The prudent course is to reserve this hypothesis for later testing in future research.

One remarkable feature of the data unmentioned before is the fact that in all the tests undertaken, never was there found a main effect! This would be noteworthy in any research area (as Campbell and Stanley, 1963, note, main effects are much more frequent in scientific research than interactions) but it is of particular interest here. Although it was a real possibility, and found by several other researchers, it did not seem to be easier to raise than to lower self-appraisal. In fact, apparently the reverse was true: the control group, which did not undergo any intervening treatments, gained 1.17 units in SEI score; the experimental groups, half of which underwent treatments prompting gain, the other half, prompting loss of self-esteem, gained only .39, on the average. Main effects for Team Success-Failure, though eminently plausible, were avoided in the present research.

Individual Difference Variables

Several individual difference variables were measured and attempts to relate them to identification and self-esteem change were intended. Since, however, self-

esteem did not change in ways that can in any way be considered identification or social comparison effects, the tests could not be performed as planned.

Some substitute (less satisfactory) methods of testing the hypotheses could be employed utilizing the variables of A-Eval, or estimated number of errors, as measures of S's introjection of M's characteristics. This was suitable in the No-Score conditions, but, generally, in the Score conditions it was undoubtedly the feedback to S about his own performance that was primarily reflected in these measures.

To summarize the individual difference hypotheses and their tenability with respect to the data:

(3) It was hypothesized that increases in strength of identification increased the magnitude of both identification and social comparison effects. In its compromised state (compromised due to the failure of SEI, that is) this hypothesis amounted to predictions of increased social comparison effects in the Score condition (i.e., a negative correlation between strength of identification and ability evaluation for Team Success, positive for Team Failure) and increased identification effects in the No Score condition (i.e., the opposite).

Using the variable of A-Eval, it appeared that in the No-Score condition there was no consistent tendency for either those with high or low strength of identification, as determined by a measure of affective identification,

to undergo vicarious self-appraisal. Instead, those with high strength of identification tended to evaluate themselves highly, independently of the performance of their classmates. Partialing out self-esteem from this relationship, on the assumption that the obtained correlation was a result of the association of each of the above variables with it, did not appreciably reduce the r. In the Score condition, where the relation of A-Eval to strength of identification could be considered a replication of Festinger, Torrey, and Willerman (1954), the replication failed. There was no tendency for those whose strength of identification was high to exhibit greater social comparison effects. It was the case, however, that Score Ss showed a smaller relationship between strength of identification and A-Eval than did No-Score Ss.

It remained unclear how to satisfactorily account for the above results.

It can be said on balance, though, that Hypothesis 3 was clearly not supported by the data. Neither social comparison effects nor identification effects were strengthened if the identificatory bond were strong.

Inasmuch as we studied but one type of identification bond (with a unique sample) it would be unwise to reject the hypothesis with regard to identification generally. The hypothesis yet deserves additional test with other types of identification than sex-role identification.

(4) Those whose initial level of self-esteem is low should be more apt to engage in identification processes than those for whom it is high. The data did not support this hypothesis. With A-Eval as a measure of vicarious self-appraisal, there was no systematic relationship of initial level of self-esteem and identification effects. With a secondary measure, estimated task score, however, there was a consistent though weak tendency for introjecting Ss to be those with high self-esteem, a result opposed to hypothesis.

Actually, this issue remains in considerable controversy, for there are a number of studies with results on either side of the question. Some find positive, some find negative correlations between self-esteem and tendency to undergo identification effects. Apparently, such concomitant conditions as the nature of subjects, the nature of identification, the nature of the task, etc., influence the course of this association. It would be interesting to conduct a study where one or more of these presumably interacting influences are varied while it is observed whether those high or those low in initial level of self-esteem are those who identify.

(5) Field-dependent Ss should exhibit greater identification effects than field-independent Ss. We expected a negative correlation for Team Success and a positive correlation for Team Failure between Field-Dependence and self-appraisal. For self-evaluation of the ability, after

partialing out a contaminating variable, the latter but not the former was obtained. Moreover, there was a tendency for the correlation to be higher for Team Failure than for Team Success, as suggested by the hypothesis. When the measure of identification effects was predicted score, however, this tendency all but disappeared.

While the hypothesis clearly can not be considered confirmed, the data are encouraging and suggestive. Field-dependence is a variable closer in theoretical substance to tendency to identify than is birth order and authoritarianism, both of which have previously been employed. The data suggest that a study whose focus is more directly on this relationship than it was in the present study might succeed in confirming that field-dependence is a predictor of disposition to identify.

(6) The final hypothesis was not directed at the question of identification vs. social comparison effects. Rather it was concerned with the relationship between initial self-esteem and subsequent self-appraisal. Hence it is but slightly relevant to the theoretical formulations in the early chapters; but it was easily studied in the context of the present experiment and of interest in its own right.

The hypothesis was that initial level of self-esteem acts as a filter on incoming information about self-worthiness, serving to bend the interpretation of the information in such a way as to uphold the existing

view. We predicted that a correlation between initial self-esteem and later measures of self-reappraisal would, for the sample as a whole, be positive.

For changes in self-esteem, the hypothesis must be rejected. The opposite tendency was found. Because of irregularities with the instrument, and other complications, this was not deemed a particularly honest test, however.

The instrument of A-Eval was a more appropriate measure of self-appraisal to test the hypothesis. For Score Ss, this was a measure of how they evaluated an ability for which they had just been tested and received feedback. The results were weak, and marginally significant, in favor of the hypothesis. We must add this finding to the growing body of literature supporting this hypothesis.

Concluding Statement

There were a number of surprising findings in this study. At least three significant findings can be considered of this type: (1) The interaction of Competition with Score on the variable of change in self-esteem; (2) The interaction of Competition with Team Success-Failure on the variable of perceived importance, where the direction of difference was the reverse of predicted; (3) The "main effect" for Score in the correlation of strength of identification with A-Eval.

The post-hoc interpretations that were offered for each of these findings were not completely satisfactory, for in each case, implications of the arguments failed to be substantiated in total by the available data. Moreover, they have no discernable pattern when taken in concert. In this respect, the study perhaps raised more questions than answered. It would be interesting to subject these unpredicted effects to further empirical test. Perhaps the interpretations proposed can be taken as propositions upon which to develop future research, as a start in answering these questions.

On balance, it can be said that none of the hypotheses were wholly supported. This was the case in spite of the fact that several of the predictions had, in previous research, obtained empirical confirmation, often under similar experimental conditions.

It is possible that since all the measures were verbal, the surprising results can be better understood by studying the unusual verbal habits of the sample. The fact of their obvious scholastic inferiority has already been mentioned. Another bit of data substantiating this assertion has not yet been presented. In the pilot test phase of this experiment, it was decided to follow Coopersmith's suggestion (personal communication) to employ Rosenberg's (1964) self-esteem measure along with his own on the basis of the fact that in "middle class schools . . . the total scores in that test consistently correlate better than .80

with the results of the SEI." In our pilot test replication the correlation was instead -.54, an utter reversal. Although stunning at the time, this inplausible discrepancy only begins to take on real significance when considered along with the other failures to replicate and unpredicted findings. It becomes not unrealistic to argue that there is something about this sample which causes them to behave, at least, on written measures, unpredictably. Here is an instance where findings, principles and hypotheses true for more typical groups (i.e., college sophomores, middle class children, etc.) do not generalize to a population studied less often. We would be wise to utilize such populations more often in social psychological research, so that we can revise our theoretical statements in order that they have greater external validity, i.e., are more universally true.

Nevertheless, when the investigator is convinced, as he is in the present case, that these same findings would not be obtained with another sample, it speaks of the anomalous character of the present sample.

We must conclude with a statement which, though overused, is particularly appropriate in the present case. In order to assess the generality of the unpredicted findings and their interpretations, in order to appraise the character of this atypical population, indeed in order to render final judgment on the hypotheses and theoretical formulations in the present paper, new

research, both straight replications with different samples and convergent validity studies, will be needed.

REFERENCES

Anderson, H. H., and Brandt, H. F. Study of motivation involving self announced goals of fifth grade children and the concept of level of aspiration. Journal of Social Psychology, 1939, 10, 209-232.

Anderson, N. H. Likeableness ratings of 555 personality-trait adjectives. Unpublished manuscript, University of California, Los Angeles, 1964.

Back, K. W. Influence through social communication. Journal of Abnormal and Social Psychology, 1951, 46, 9-23.

Bandura, A. Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.) Advances in Experimental Social Psychology, Vol. 2. New York: Academic Press, 1965.

Bandura, A. Influence of model's reinforcement contingencies on the acquisition of imitative responses. Journal of Personality and Social Psychology, 1965, 1, 589-595.

Bandura, A. Principles of behavior modification. New York: Holt, Rinehart and Winston, Inc., 1969.

Bandura, A., Grusec, J. E., and Menlove, F. L. Observational learning as a function of symbolization and incentive set. Child Development, 1966, 37, 499-506 (a).

Bandura, A., Grusec, J. E., and Menlove, F. L. Some social determinants of self-monitoring reinforcement systems. Journal of Personality and Social Psychology, 1967, 5, 449-455. (b)

Bandura, A., and Harris, M. D. Modification of syntactic style. Journal of Experimental Child Psychology, 1966, 4, 341-352.

Bandura, A., and Huston, A. C. Identification as a process of incidental learning. Journal of Abnormal and Social Psychology, 1961, 63, 311-318.

Bandura, A., and Kupers, C. J. The transmission of patterns of self-reinforcement through modeling. Journal of Abnormal and Social Psychology, 1964, 69, 1-9.

Bandura, A., and McDonald, F. J. The influence of social reinforcement and the behavior of models in shaping children's moral judgements. Journal of Abnormal and Social Psychology, 1963, 67, 274-281.

Bandura, A., and Menlove, F. L. Factors determining vicarious extinction of avoidance behavior through symbolic modeling. Journal of Personality and Social Psychology, 1968, 8, 99-108.

Bandura, A., and Mischel, W. The influence of models in modifying delay of gratification patterns. Journal of Personality and Social Psychology, 1965, 2, 698-705.

Bandura, A., and Rosenthal, T. L. Vicarious classical conditioning as a function of arousal level. Journal of Personality and Social Psychology, 1963, 3, 54-62.

Bandura, A., and Ross, D., and Ross, S. A. Imitation of film mediated aggressive models. Journal of Abnormal and Social Psychology, 1963, 66, 3-11. (a)

Bandura, A., Ross, D., and Ross, S. A. A comparative test of the status envy, social power, and secondary reinforcement theories of identificatory learning. Journal of Abnormal and Social Psychology, 1963, 67, 527-534. (b)

Bandura, A., and Walters, R. H. Social learning and personality development. New York: Holt, Rinehart, and Winston, Inc., 1963.

Bandura, A., and Whalen, C. K. The influence of antecedent reinforcement and divergent modelling cues on patterns of self-reward. Journal of Personality and Social Psychology, 1966, 3, 373-382.

Baxter, J. C., Lerner, M. J., and Miller, J. S. Identification as a function of the reinforcing quality of the model and the socialization background of the subject. Journal of Personality and Social Psychology, 1965, 2, 692-697.

Berger, S. M. Conditioning through vicarious instigation. Psychological Review, 1962, 69, 450-466.

Bettleheim, B. Individual and mass behavior in extreme situations. Journal of Abnormal and Social Psychology, 1943, 38, 417-452.

Braver, S. L. Self-concept and status. Unpublished manuscript, Wayne State University, Detroit, 1964.

Brehm, J., and Festinger, L. Pressures toward uniformity of performance in group. Human Relations, 1957, 10, 85-91.

Brenner, M. The effects of love on interference of recall. In preparation, University of Michigan, Ann Arbor, 1970.

Bronfenbrenner, U. The study of identification through the interpersonal perception. In R. Tagiuri, and L. Petrullo (Eds.) Person perception and interpersonal behavior. Stanford, California: Stanford University Press, 1958.

Bronfenbrenner, U. Freudian theories of identification and their derivatives. Child Development, 1960, 31, 15-40.

Brown, R. Social psychology. New York: The Free Press, 1965.

Burnstein, E., Stotland, E., and Zander, A. Similarity to a model and self-evaluation. Journal of Abnormal and Social Psychology, 1961, 62, 257-264.

Byrne, D. The Repression-Sensitization Scale: Rationale in reliability, and validity. Journal of Personality, 1961, 29, 334-349.

Byrne, D. Repression-sensitization as a dimension of personality. In B. A. Maher (Ed.) Progress in experimental personality research. New York: Academic Press, 1964.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1963.

Carlson, R. Identification and personality structure in pre-adolescents. Journal of Abnormal and Social Psychology, 1963, 67, 566-573.

Cetlin, I. G. Persistence, defensive behavior, and self-evaluation as a function of early and late anticipation of failure. Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1964.

Chapman, D. W., and Volkmann, J. A social determinant of the level of aspiration. Journal of Abnormal and Social Psychology, 1939, 34, 225-238.

Cooley, C. H. Human nature and the social order. New York: Charles Scribner and Sons, 1902.

Coopersmith, S. The antecedents of self-esteem. San Francisco: Freeman Press, 1968.

Couch, C. J., and Murray, J. S. Significant others and evaluation. Sociometry, 1964, 27, 502-509.

Cronbach, L. J. and Furby L. How we should measure "change" - or should we? Unpublished manuscript, Stanford University, Stanford, California, 1969.

Cutick, R. A. Self-evaluation of capacities as a function of self-esteem and the characteristics of a model. Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1962.

Davis, J. A. The campus as a frog-pond: An application of the theory of relative deprivation to career decisions of college men. American Journal of Sociology, 1966, 72, 17-31.

Diggory, J. C. Self evaluation: Concepts and studies. New York: Wiley and Sons, 1966.

Dittes, J. E. Effects of changes in self-esteem upon impulsiveness and deliberation in making judgements. Journal of Abnormal and Social Psychology, 1959, 59, 77-82.

Dreyer, A. S. Aspiration behavior as influenced by expectation and group comparison. Human Relations, 1954, 7, 175-190.

Ferguson, E. D. Effect of sibling competition and alliance on level of aspiration, expectation and performance. Journal of Abnormal and Social Psychology, 1958, 56, 213-223.

Festinger, L. Wish, expectation and group standards as factors influencing level of aspiration. Journal of Abnormal and Social Psychology, 1942, 37, 184-200.

Festinger, L. A theory of social comparison processes. Human Relations, 1954, 7, 117-140.

Festinger, L., Torrey, J., and Willerman, B. Self-evaluation as a function of attraction to the group. Human Relations, 1954, 7, 161-174.

Fishbein, M., Raven, B. H., and Hunter, R. Social comparison and dissonance reduction in self-evaluation. Journal of Abnormal and Social Psychology, 1963, 67, 491-501.

French, J. R. P. The social environment and mental health. Journal of Social Issues, 1963, 19, 39-56.

Freud, S. New introductory lectures on psychoanalysis. New York: Norton Press, 1933.

Freud, S. An outline of psychoanalysis. New York: Norton Press, 1949. (a)

Freud, S. The passing of the Oedipus-complex. In Collected Papers, Vol. 2. London: Hogarth Press, 1949. (b)

Gardner, J. W. Level of aspiration in response to a prearranged sequence of scores. Journal of Experimental Psychology, 1939, 25, 601-621.

Gerard, H. B. Some determinants of self-evaluation. Journal of Abnormal and Social Psychology, 1961, 62, 288-293.

Gergen, K. J., and Morse, S. J. Social comparison self-consistency and the presentation of self. Paper presented at the American Psychological Association Meeting, 1967.

Goodenough, D. R., and Eagle, C. J. A modification of the embedded figures test for use with young children. Journal of Genetic Psychology, 1963, 103, 67-94.

Guilford, J. P. Psychometric methods. New York: McGraw-Hill, 1936.

Haas, H. I., and Maehr, M. L. Two experiments on the concept of self and the reaction of others. Journal of Personality and Social Psychology, 1965, 1, 100-105.

Hakmiller, K. Threat as a determinant of downward comparison. Journal of Experimental Social Psychology, Supplement 1, 1966, 32-39.

Hansche, J., and Gilchrist, J. C. Three determinants of the level of aspiration. Journal of Abnormal and Social Psychology, 1956, 53, 136-137.

Heider, F. The psychology of interpersonal relations. New York: Wiley and Sons, Inc., 1958.

Helper, M. M. Learning theory and the self-concept. Journal of Abnormal and Social Psychology, 1955, 51, 184-194.

Helper, M. M. Parental evaluations of children and children's self-evaluation. Journal of Abnormal and Social Psychology, 1958, 56, 190-194.

Hicks, J. M. The influence of group flattery upon self-evaluation. Journal of Social Psychology, 1962, 58, 147-141.

Hilgard, E. R., Sait, E. M., and Magaret, G. A. Level of aspiration as affected by relative standing in an experimental social group. Journal of Experimental Psychology, 1940, 27, 411-421.

Hoffman, P. J., Festinger L., and Lawrence, D. H. Tendencies toward group comparibility in competitive bargaining. Human Relations, 1954, 7, 141-159.

Isen, A.M. Success, failure, attention and reaction to others: The warm glow of success. Journal of Personality and Social Psychology, 1970, 15, 294-301.

Izard, C. E. Personality correlates of sociometric status. Journal of Applied Psychology, 1959, 43, 89-93.

James, W. The principles of psychology, Vol. 2. New York: Holt, 1890.

Johnston, W. A. Self-evaluation in a simulated team. Psychonomic Science, 1966, 6, 261-262.

Johnston, W. A. Individual performance and self-evaluation in a simulated team. Organizational Behavior and Human Performance, 1967, 2, 309-328.

Johnston, W. A., and Nawrocki, L. H. The effects of simulated team feedback on the performance of good and poor trackers. Air Force Office of Scientific Research, Technical Report, 66-2541, 1966.

Johnston, W. A., and Nawrocki, L. H. Effects of simulated social feedback on individual tracking performance. Journal of Applied Psychology, 1967, 51, 145-151.

Jones, E. E., and Gerard, H. B. Foundations of social psychology. New York: Wiley and Sons, Inc., 1967.

Kagan, J. The concept of identification. Psychological Review, 1958, 65, 296-305.

Kagan, J., and Phillips, W. Measurement of identification: A methodological note. Journal of Abnormal and Social Psychology, 1964, 69, 442-447.

Kaufman, B. H. The use of embedded figures tests. Unpublished manuscript, Rutgers, The State University, New Brunswick, New Jersey, 1965.

Kuhn, M. H., and McPartland, T. S. An empirical investigation of self-attitudes. American Sociological Review, 1954, 19, 68-76.

Latane, B., Studies in social comparison: Introduction and overview. Journal of Experimental Social Psychology Supplement 1, 1966, 1-5.

Lazowick, L. M. On the nature of identification. Journal of Abnormal and Social Psychology, 1955, 51, 175-184.

Lefcourt, H. M., and Ladwig, G. W. The effects of reference group upon Negroes' task persistence in a biracial competitive game. Journal of Personality and Social Psychology, 1965, 1, 668-671.

Lewin, K., Dembo, T., Festinger, L. and Sears, P. S. Level of aspiration. In J. McV. Hunt (Ed.) Personality and the behavior disorders. New York: Ronald Press, 1944.

Lynn, D. B. Parental and sex-role identification: A theoretical formulation. Berkely: McCutchan Press, 1969.

Maccoby, E. E. Role-taking in childhood and its consequences for social learning. Child Development, 1959, 30, 239-252.

Maehr, M. L., Mensing, J., and Nafzger, S. Concept of self and the reactions of others. Sociometry, 1962, 25, 353-357.

Maliver, B. L. Anti-Negro bias among Negro college students. Journal of Personality and Social Psychology, 1965, 2, 770-775.

Manis, M. Social interaction and the self concept. Journal of Abnormal and Social Psychology, 1955, 51, 362-370.

Mead, G. H. Mind, self, and society. Chicago: University of Chicago Press, 1934.

Miller, D. R. The study of social relationships: Situation, identity, and social interaction. In S. Koch (Ed.) Psychology: A Study of a Science, Vol. 5, New York: McGraw-Hill, 1963.

Miyamoto, S. F., and Dornbusch, S. M. A test of interactionist hypotheses of self-conception. American Journal of Sociology, 1956, 61, 399-403.

Mowrer, O. H. Identification: A link between learning theory and psychotherapy. In Learning theory and personality dynamics. New York: Ronald Press, 1950.

Mussen, P. Some antecedents and consequents of masculinity sex-typing in adolescent boys. Psychological Monographs, 1961, 75, Whole number 506.

Mussen, D. Early socialization, learning, and identification. In T. Newcomb (Ed.) New directions in psychology, Vol. 3. New York: Holt, Rinehart, and Winston, 1967.

Mussen, P., and Distler, L. Masculinity identification and father-son relationships. Journal of Abnormal and Social Psychology, 1963, 59, 350-356.

Nunnally, J. Psychometric theory. New York: McGraw-Hill, 1967.

Parsons, T., and Shils, E. A. (Eds.) Toward a general theory of action. Cambridge, Massachussetts: Harvard University Press, 1951.

Payne, D. E., and Mussen, P. Parent-child relations and father-identification among adolescent boys. Journal of Abnormal and Social Psychology, 1956, 52, 358-362.

Pettigrew, T. F. Social evaluation theory: Convergences and applications. In D. Levine (Ed.) Nebraska Symposium on Motivation. Lincoln, Nebraska: University of Nebraska Press, 1967.

Radloff, R. Social comparison and ability evaluation. Journal of Experimental Social Psychology, Supplement 1, 1966, 6-26.

Raven, B. H., and Fishbein, M. Social referents and self-evaluation in examinations. Journal of Social Psychology, 1965, 65, 89-99.

Raven, B. H., and Rietsema, J. The effects of varied clarity of group goal and group path upon the individual and his relation to his group. Human Relations, 1957, 10, 29-46.

Rosenberg, M. Parental interest and children's self-conception. Sociometry, 1963, 26, 35-49.

Rosenberg, M. Society and the adolescent self-image. Princeton: Princeton University Press, 1965.

Rothman, D. Z. The effects of success and failure experiences in normal, "neurotic," and schizophrenic populations. Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1963.

Sanford, N. The dynamics of identification. Psychological Review, 1955, 62, 106-118.

Sarnoff, I. Identification with the aggressor: Some personality correlates of anti-Semitism among Jews. Journal of Personality, 1951, 20, 199-218.

Schachter, S. The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.) Advances in Experimental Social Psychology, Vol 1. New York: Academic Press, 1964.

Sears, R. R. Identification as a form of behavioral development. In D. Harris (Ed.) The concept of development. Minneapolis: University of Minnesota Press, 1957.

Sharma, S. L. Some personality correlates of changes in self-esteem under conditions of stress and support. Journal of Educational Psychology, 1956, 14, 154-165.

Singer, J. E. Social comparison: Progress and issues. Journal of Experimental Social Psychology, Supplement 1, 103-110.

Stoke, S. M. An inquiry into the concept of identification. In W. E. Martin and C. B. Stendler (Eds.) Readings in child development. New York: Harcourt, Brace and Co., 1954.

Stotland, E. Exploratory investigations of empathy. In L. Berkowitz (Ed.) Advances in Experimental Social Psychology, Vol. 4. New York: Academic Press, 1969.

Stotland, E., and Cottrell, N. B. Self-esteem, group interaction and group influence on performance. Journal of Personality, 1961, 29, 273-284.

Stotland, E., and Cottrell, N. B. Similarity of performance as influenced by interaction, self-esteem, and birth-order. Journal of Abnormal and Social Psychology, 1962, 64, 183-191.

Stotland, E., Cottrell, N. B. and Laing, G. Group interaction and perceived similarity of members. Journal of Abnormal and Social Psychology, 1960, 61, 335-340.

Stotland, E., and Dunn, R. E. Identification, "opposite-ness," authoritarianism, self-esteem, and birth order. Psychological Monographs, 1962, 76, Whole number 528.

Stotland, E., and Dunn, R. E. Empathy, self-esteem and birth-order. Journal of Abnormal and Social Psychology, 1963, 66, 532-540.

Stotland, E., and Hillmer, M. L. Identification, authoritarianism, defensiveness and self-esteem. Journal of Abnormal and Social Psychology, 1962, 64, 334-342.

Stotland, E., Katz, D., and Patchen, M. The reduction of prejudice through the arousal of self-insight. Journal of Personality, 1959, 27, 507-531.

Stotland, E., and Patchen, M. Identification and changes in prejudice and in authoritarianism. Journal of Abnormal and Social Psychology, 1961, 62, 265-274.

Stotland, E., Thorley, S., Thomas, E., Cohen, A. R., and Zander, A. The effects of group expectations and self-esteem upon self-evaluation. Journal of Abnormal and Social Psychology, 1957, 54, 55-63.

Stotland, E., and Walsh, J. A. Birth order and an experimental study of empathy. Journal of Abnormal and Social Psychology, 1963, 66, 610-614.

Stotland, E., Zander, A., and Natsoulas, T. Generalization of interpersonal similarity. Journal of Abnormal and Social Psychology, 1961, 62, 250-256.

Thornton, D. A., and Arrowood, A. J. Self-evaluation, self-enhancement, and the locus of social comparison. Journal of Experimental Social Psychology, Supplement 1, 1966, 40-48.

Thurstone, L. L. A factorial study of perception. Chicago: University of Chicago Press, 1944.

Videbeck, R. Self-conception and the reactions of others. Sociometry, 1960, 22, 351-359.

Walker, E. E. Racial and national identifications. Journal of Negro Education, 1965, 34, 17-20.

Wheeler, L. Motivation as a determinant of upward comparison. Journal of Experimental Social Psychology, Supplement 1, 1966, 27-31.

Whiting, J. W. M. Resource mediation and learning by identification. In I. Iscoe and H. W. Stevenson (Eds.) Personality development in children. Austin, Texas: University of Texas Press, 1960.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Witkin, H. A., Dyk, R. B., Paterson, H. F., Goodenough, D. R., and Karp, S. A. Psychological differentiation. New York: Wiley and Sons, Inc., 1962.

Witkin, H. A., Lewis, A. B., Hertzman, M., Machover, M., Meissner, P. B., and Wapner, S. Personality through perception: An experimental and clinical study. New York: Harper and Row, Inc., 1954.

Wylie, R. C. The self-concept. Lincoln: The University of Nebraska Press, 1961.

Zander, A. Group membership and individual security. Human Relations, 1958, 11, 99-112.

Zander, A., Stotland, E., and Wolfe, P. Unity of the group, identification with the group, and the self-esteem of members. Journal of Personality, 1960, 28, 463-478.

APPENDIX A.1

LETTER TO PARENTS OF SUBJECTS

The following letter was sent to parents of all prospective subjects:

Dear parent:

Your child's class has been selected by researchers at the University of Michigan to participate in a study during the month of February, 1970. The research is of a scientific nature, exploring conditions affecting children's self-esteem. All information to be gained will be confidential, and will be used for research purposes exclusively; that is, conclusions will be about children generally, rather than about specific children. This study will involve approximately two hours of class time at most. This project has the approval of your child's classroom teacher and the Building Principal.

If you do not wish your son or daughter to participate in this study, please check the box below and return this slip to school.

Sincerely yours,

Sanford L. Braver

Melvin Manis, Ph.D.

☐ I do not wish my child to participate in this research project.

_____

(Parent's Signature)

APPENDIX A.2

MEMO TO TEACHERS OF SUBJECTS

The following memo was sent to the teachers of subjects' classrooms:

MEMO TO: Fifth and Sixth Grade Teachers, Mt. Pleasant and Cory Schools

FROM: Sanford Braver

RE: Research Project involving your students

In the interests of insuring the smoothest, least disruptive, most valid opportunity for research results, I make the following suggestions.

BEFORE PART I

1. Speak as little as possible about the study to your classes. Answer your students' questions, if any, as non-commitally as possible.
2. Make a list of those students whose self-esteem, in your judgement, is very fragile. We will be especially careful with these. Keep in mind also, and inform me of, any student who, because of reading or other difficulties, are likely not to comprehend instructions, etc.
3. Remember the names of students whose parents wished them not to participate.

DURING PART I

1. Assist me in keeping order in the class.
2. Assist me in clarifying questions. If his question is other than what a word means, or similar easily answered questions, it is best to let me answer the question, however.
3. Give students whose parents wished them not to participate special work. They may be sent to the library, etc., if you prefer.
4. Give me the names on lists you have kept.

BETWEEN PARTS I AND II

1. Be non-commital about questions. The most frequently asked question will probably be immediately after I leave: "Is he coming back to give us our scores?" Feign ignorance, be vague. If they know too much it may invalidate the conclusions.

DURING PART II

This session, a few days after the first, will involve only the boys. I will come in and call usually 4 of them at a time, each group for a half-hour or less. They will accompany me to another room.

1. Most importantly, try to prevent kids who have already been tested from speaking to others who have not. I ask them to keep my secret, and most probably will; use your judgement on how to do this. The best approach probably is to keep them busy enough so they won't have time to compare notes with one another. Also, prevent people who have already been tested from speaking to others who have also been tested.
2. Be aware of the things that <u>are</u> said regarding the study by kids who have been tested, and report them to me. I am particularly concerned with inconsistent stories about what happened. If two kids disagree about what I told them, again feign ignorance, and report it to me.

AFTER PART II

1. Again, be aware of disagreements that might come to light after I've completed Part II. Answer that you don't know, but maybe I or they made a mistake.
2. A day or two after I finish Part II, you might announce that I've scored the <u>girls'</u> tests and they, <u>too</u>, did very well on the test!

I'll be very happy to answer any questions you might have.

APPENDIX A.3

ORAL AND WRITTEN INSTRUCTIONS AND INSTRUMENT FOR INITIAL MEASUREMENT OF SELF-ESTEEM

The following are the oral instructions for the initial measurement of self-esteem. The instrument appears on page 190.

> Today I am going to give your class some forms to fill out. Don't open them until I tell you to. I'll need one helper to pass out these forms to the boys, another to pass out girls' forms and a third to pass out pencils to everyone. Use the red pencils to mark your answers, to all the questions.
>
> When you get your form and your pencil, write your name and school at the top. Where it says class put your teachers' name. Now I'll read the instructions for the first page aloud while you follow.

Read instructions.

> Remember, these are just your opinion. If you can't read a word or don't understand the instructions raise your hand, and you will be helped. When you have finished, don't turn the page but put down your pencil and wait for me to give you your next instructions. If you have to erase an answer here or on any other pages, make sure you erase it clearly enough so that I'll know which is your answer.

Name_____ School_____

Class_____ Boy or Girl? Boy

Please mark each statement in the following way:
If the statement describes how you usually feel, put a check (✓) in the column "LIKE ME."
If the statement does not describe how you usually feel, put a check (✓) in the column "UNLIKE ME."
There are no right or wrong answers.

	LIKE ME	UNLIKE ME
Example: I'm a hard worker		
1. I spend a lot of time daydreaming.		
2. I'm pretty sure of myself.		
3. I often wish I were someone else.		
4. I'm easy to like.		
5. I can make up my mind without too much trouble.		
6. I get upset easily at home.		
7. Someone always has to tell me what to do.		
8. It takes me a long time to get used to anything new.		
9. I'm often sorry for the things I do.		
10. I give in very easily.		
11. I'm pretty happy.		
12. I would rather play with children younger than me.		
13. I like to be called on in class.		
14. I understand myself.		
15. Things are all mixed up in my life.		
16. Kids usually follow my ideas.		
17. I'm not as nice looking as most people.		
18. If I have something to say, I usually say it.		
19. Kids pick on me very often.		
20. My parents understand me.		
21. I don't care what happens to me.		
22. I'm a failure.		
23. Most people are better liked than I am.		
24. I often get discouraged in school.		

APPENDIX A.4

ORAL AND WRITTEN INSTRUCTIONS AND INSTRUMENT FOR MEASUREMENT OF COGNITIVE IDENTIFICATION

The following are the oral instructions for the measure of cognitive identification. The instrument appears on pages 192 and 193.

> Now we are ready to start the next page. Turn the page and read the instructions to yourself while I read them out loud. If you are a girl what you read will be a little different than what I read out loud. Whenever I say "boy" your instructions will say "girl". Also the names will be different. Whenever I say the name Mel, your form will say Peg and whenever I say Mark, your form will say Mary; otherwise the instructions will be the same, for boys and girls.

Read instructions.

> If you understand the instructions now you may turn the page and start to answer. Remember that the answers are just your opinion. There are no right or wrong answers. If you have any questions about the instructions or if you can't read or understand what any of the words mean, raise your hand and you will be helped. Circle only one answer on each line. When you have finished this page, put down your pencil and wait quietly.

OPINIONS ABOUT BOYS AND GIRLS

INSTRUCTIONS

On the following page we will be talking about different boys. We will describe boys two at a time, one on each side of the page. You should read about each boy we describe very carefully, and think about whether you agree with him or not. Beneath the descriptions are a row of statements. You should circle the statement that best identifies how you feel.

Here is an example to show you how you should answer:

MEL			MARK	
Mel thinks that arithmetic is harder than spelling.			Mark thinks that spelling is harder than arithmetic.	
I agree completely with Mel	I agree more with Mel than with Mark	I agree as much with Mel as with Mark	(I agree more with Mark than with Mel)	I agree completely with Mark

Which you should circle depends upon how you feel about arithmetic and spelling. If you thought arithmetic was always very hard, and spelling was easy and fun, you would circle the first statement, "I agree completely with Mel". If you thought that sometimes arithmetic was easier than spelling, but usually arithmetic was harder, you would circle the second statement, "I agree more with Mel than with Mark." Circle the first statement only when you agree completely with that person. If you thought that arithmetic was about as hard as spelling, you would circle the third statement, "I agree as much with Mel as with Mark."

The person who answered the example thought that spelling was usually harder than arithmetic, but not always, so he agreed more with Mark than with Mel.

Answer the questions on the following page in this way. Turn the page and begin.

JIM

Jim thinks that boys and girls are very much the same. He thinks boys mostly like to do just about the same things that girls like to do.

JACK

Jack thinks that boys and girls are very different. He thinks that girls like different things and act differently than boys do.

I agree completely with Jim	I agree more with Jim than with Jack	I agree as much with Jim as with Jack	I agree more with Jack than with Jim	I agree completely with Jack

BILL

Bill thinks that boys are all pretty different. He thinks every boy has a lot of likes and dislikes that other boys don't share.

BOB

Bob thinks that most boys are a lot alike. He thinks most boys have the same likes and dislikes as every other boy.

I agree completely with Bill	I agree more with Bill than with Bob	I agree as much with Bob as with Bill	I agree more with Bob than with Bill	I agree completely with Bob

DON

Don is not too much like the other boys in his class. He likes a lot of things and does a lot of things that are different from the things other boys his age like and do.

DAVE

Dave is a lot like most of the other boys in his class. He likes and does just about the same things that other boys his age like and do.

I am completely like Don	I am more like Don than Dave	I am as much like Don as I am like Dave	I am more like Dave than Don	I am completely like Dave

TODD

Todd thinks that boys should stick together against girls. He thinks that boys who play or talk to girls are sissies. He hardly ever plays or talks to girls in his class.

TIM

Tim thinks that boys should play with girls. He thinks that girls are often as much fun to play with as boys are. He very often plays and talks to girls in his class.

I am completely like Todd	I am more like Todd than Tim	I am as much like Tim as I am like Todd	I am more like Tim than Todd	I am completely like Tim

APPENDIX A.5

ORAL AND WRITTEN INSTRUCTIONS AND INSTRUMENT FOR MEASUREMENT OF AFFECTIVE IDENTIFICATION

The following are the oral instructions for the measure of affective identification. The instrument appears on page 195.

> Now turn the page and read the instructions for the next page. If you understand the instructions go ahead and mark your answers. When you have finished put down your pencils and wait. Again there are no right or wrong answers on this part; just give your opinion. If you don't understand what you are to do, raise your hand and you will be helped.

MORE OPINIONS ABOUT BOYS AND GIRLS

INSTRUCTIONS: On this page there is a list of words that describe people. On one side of each word there is the word BOYS; on the other side of the word there is written GIRLS. For each word in the middle think of whether it better describes boys or girls. For example, if you thought that girls were more often "pretty" than boys you would circle GIRLS, as is done below.

(GIRLS)	pretty	BOYS

You might also think that the word "strong" describes boys better than girls, since boys are more often strong than girls. Then you would circle BOYS, as below:

GIRLS	strong	(BOYS)

Now continue down the list, circling either BOYS or GIRLS for each word in the center. There are no right or wrong answers; just give your opinion.

GIRLS	lazy	BOYS
GIRLS	bossy	BOYS
GIRLS	friendly	BOYS
GIRLS	honest	BOYS
GIRLS	interesting	BOYS
GIRLS	grouchy	BOYS
GIRLS	nosey	BOYS
GIRLS	smart	BOYS
GIRLS	active	BOYS
GIRLS	selfish	BOYS

APPENDIX A.6

ORAL AND WRITTEN INSTRUCTIONS AND INSTRUMENT FOR MEASUREMENT OF FIELD-DEPENDENCE

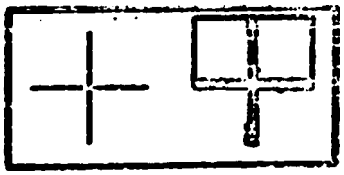
The following are the oral instructions for the measure of Field-Dependence. The instrument appears on pages 197, 198, and 199.

> Now turn the page and read the instructions on the next page. Try to work the sample problem. When you have finished reading the instructions, look up; don't turn the page. (Work sample problems on board.) Now the problems on the next two pages are timed--you will have a time limit, so you should try to work as many problems correctly as you can before I say stop. Probably none of you will finish all the problems before I stop you, so don't be surprised if you haven't finished. If you do finish, though, raise your hand as soon as you do. If you are having a lot of trouble with one problem, skip it and go on to the next. Go as quickly as possible. Ok ready! Turn the page and begin. (After 5 minutes) Whether you have finished or not go on to the next page right now. Again work as fast as you can. (After 2½ minutes) Stop! Time is up. Put down your pencils.

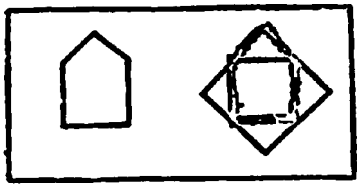
HIDING PICTURES

INSTRUCTIONS

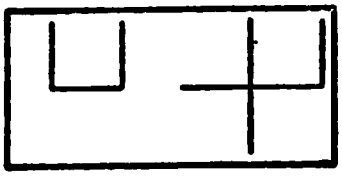
There are two drawings in the box. The second figure is really the same as the first except that some more lines have been drawn. So the first figure is contained in the second. The second figure has been marked to show this. You can see that the part of the second figure that has been marked is exactly the same as the first figure.



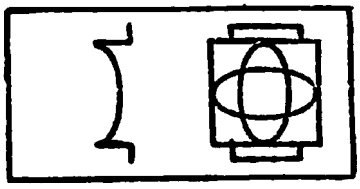
In the pair of drawings in this box, the part of the second figure which is the same as the first figure has again been marked. This is your job. In each pair, the second figure contains the first figure. You should try to find it and mark it as quickly as you can. Don't try to make your markings overly neat or accurate. Just trace the first figure clearly enough so that we can be sure you saw it. Remember that the first figure can be found someplace in the second in exactly the same size and position as it originally appeared.



Here is a sample to try to do yourself.



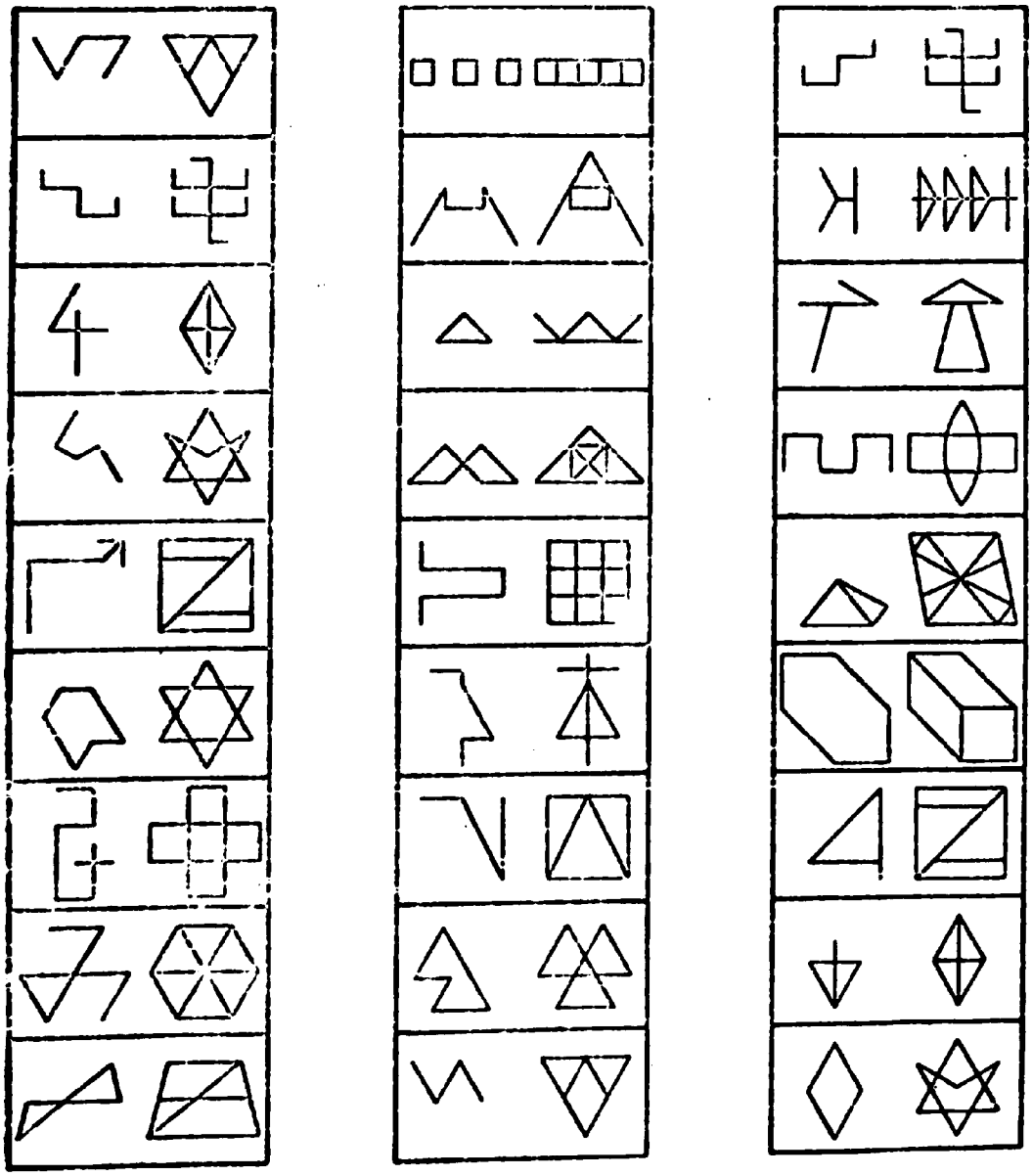
Here is another sample to mark. After you've tried it the examiner will show you the correct answer on the blackboard. Then on the next page there will be a set of such hiding pictures for you to find and mark as quickly as you can.



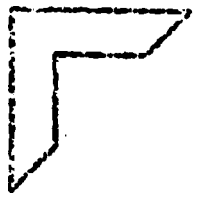
DON'T TURN THE PAGE UNTIL YOU ARE TOLD TO DO SO!

PART I

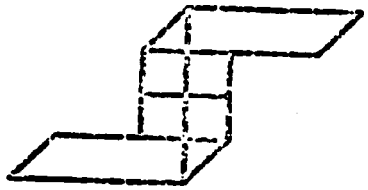
In each pair of figures below, mark that part of the second figure which is the same as the first.

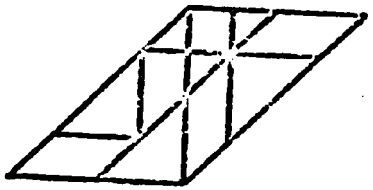


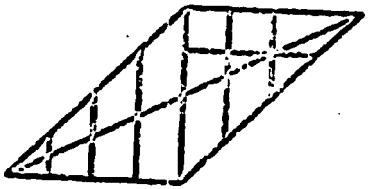
PART II

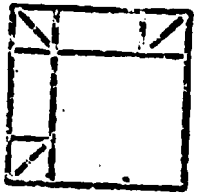


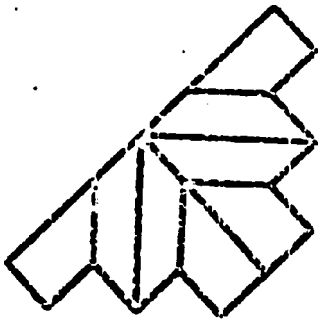
Look at the above figure.
It is contained in each of the drawings below. Find it in each drawing and then mark it. Mark only one figure in each drawing.

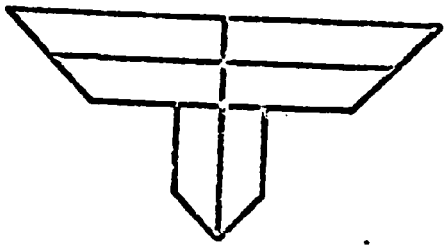


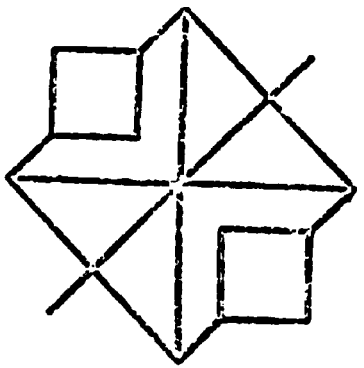












APPENDIX A.7

ORAL AND WRITTEN INSTRUCTIONS AND INSTRUMENT FOR COMPLETION OF BOGUS TASK

The following are the oral instructions for the completion of the bogus task. The instrument appears on pages 201, 202, 203, 204, 205, and 206.

> Now read the instructions on the next page. There will not be a time limit on the rest of the test and you will be allowed to work til you finish. On these problems, do not use your pencils or fingers to help make measurements so that you can answer. Answer just by looking with your eyes at the problems. If you're not sure you understand the instructions, for any page, raise your hand and you will be helped. Continue working until you have finished the test. When you have finished close your test booklet, put down your pencils and rest.

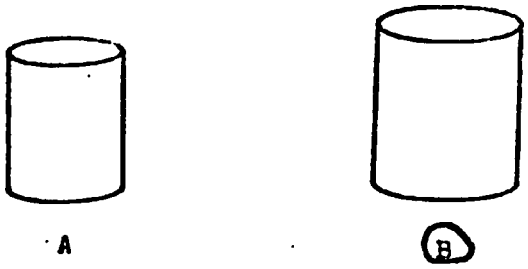
SIZES AND SHAPES

INSTRUCTIONS: All of the following questions have two possible answers, A or B. Read the instructions at the top of each page carefully, and study the example which has the correct answer circled. If you still don't understand the instructions, raise your hand and you will be helped. Then for each problem on the page you should try to figure out whether A or B is the better answer, JUST BY LOOKING AT THE PICTURES, and circle the correct answer. If you are not sure of an answer (many of them are quite hard) take the best guess you can; you may be right!

PART I: Water Glasses

Look at the following pairs of water glasses. Which one do you think holds the most water, A or B of each pair? Circle the letter of the glass which holds more water.

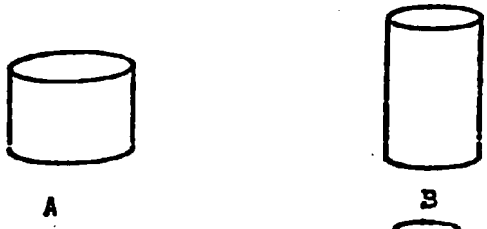
Example:



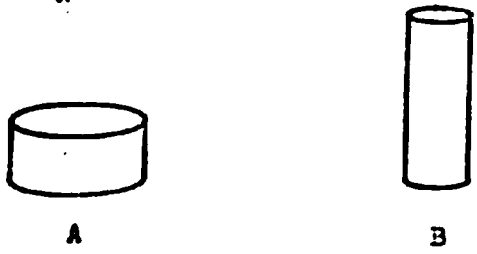
Problem 1:



Problem 2:



Problem 3:



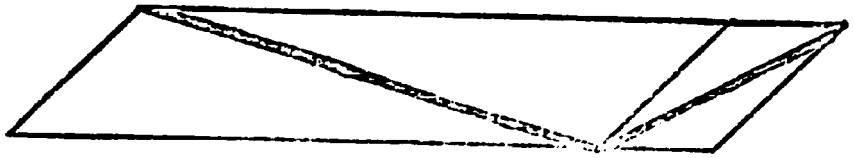
Problem 4:



PART II: Diagonal lines

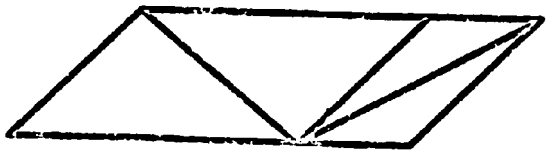
Look at the following figures. Each has two lines going from the bottom to one of the corners at the top. These are called diagonals, and they are marked in the example. Which diagonal do you think is longer? Circle the letter that lies below the diagonal line you think is longer. In the example, A is circled because A lies below the diagonal that is longer.

Example:



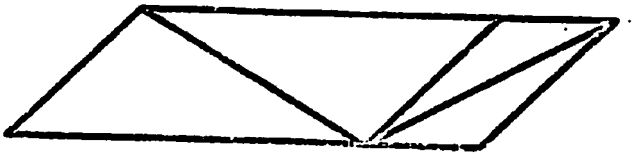
(A) B

Problem 1:



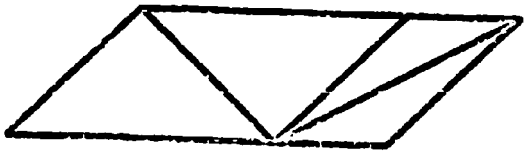
A B

Problem 2:



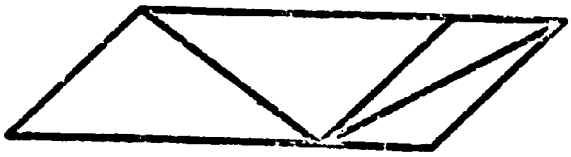
A B

Problem 3:



A B

Problem 4:

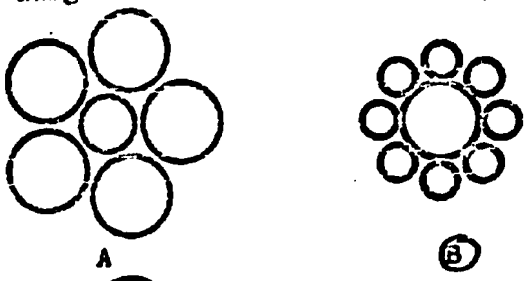


A B

PART III: Circles

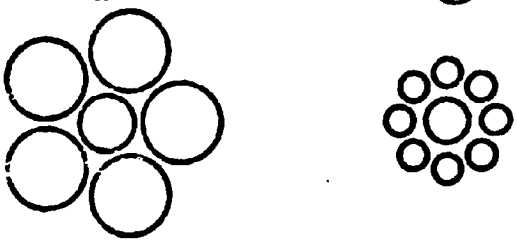
Each of the problems on this page has two groups of circles. The middle circle in each group is the important one. Which center circle is the largest? Circle the letter which lies below the largest middle circle. In the example, the largest middle circle lies above B, so B is circled.

Example:



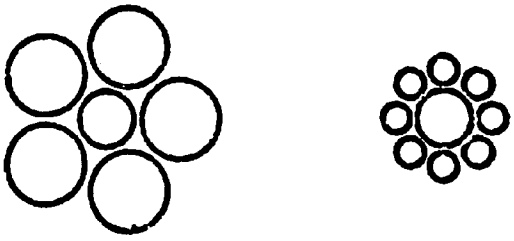
A (B)

Problem 1:



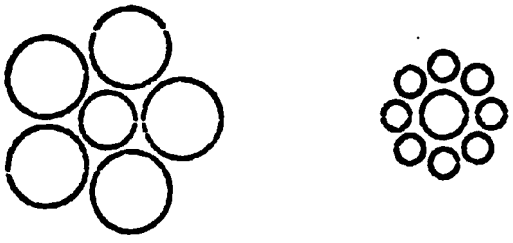
A B

Problem 2:



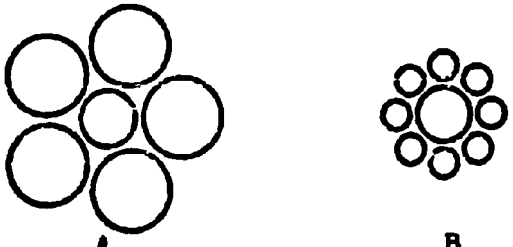
A B

Problem 3:



A B

Problem 4:

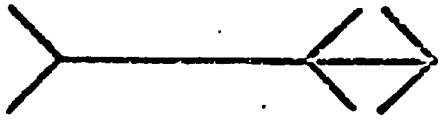


A B

PART IV: Arrows

Look at the arrows on this page. You will see that each has an extra arrowhead in the middle that divides the arrow into two parts, A & B Which of the two parts is longer? Circle the letter that lies below the longer part of the arrow.

Example:



(A) B

Problem 1:



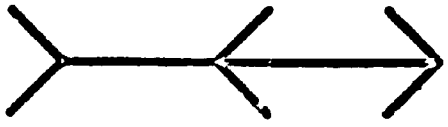
A B

Problem 2:



A B

Problem 3:



A B

Problem 4:

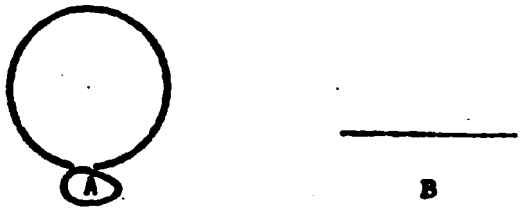


A B

PART V: Lines

In the problems below there are two lines. Line A has been bent around until the ends almost touch, so that it almost forms a circle. Line B is a straight line. If line A were straightened out into a straight line again, which would be longer, line A or line B? In the example, A is circled because line A, if it were pulled straight, would be longer than line B.

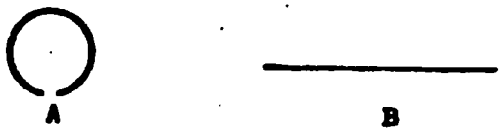
Example:



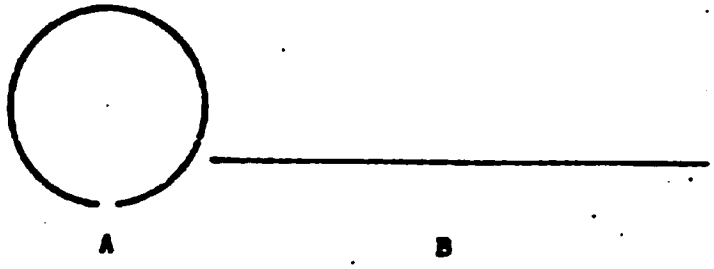
Problem 1:



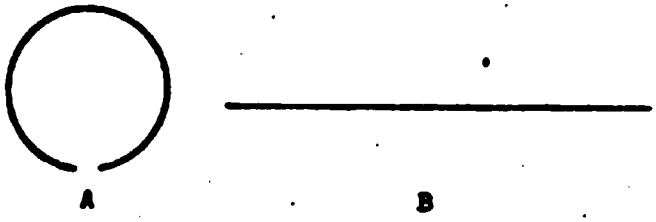
Problem 2:



Problem 3:



Problem 4:



On the test you just finished taking, there were 5 pages and 20 problems in all where you answered A or B. If you had to make a guess, how many do you think you probably got wrong? Make your best guess, then write the number you think you got wrong out of 20 below, on the line.

_____ wrong.

APPENDIX A.8

INSTRUMENT FOR FINAL MEASUREMENT OF SELF-ESTEEM

On page 208 occurs the instrument for the final measurement of self-esteem. The oral instructions appear in the text.

Name_____School_____

Class_____Boy or Girl?_____

Please mark each statement in the following way:

If the statement describes how you usually feel, put a check (✓) in the column "LIKE ME."
If the statement does not describe how you usually feel, put a check (✓) in the column "UNLIKE ME."
There are no right or wrong answers.

	LIKE ME	UNLIKE ME
Example: I'm a hard worker_____		
1. My parents and I have a lot of fun together._____		
2. I find it very hard to talk in front of the class._____		
3. I wish I were younger._____		
4. I'm a lot of fun to be with._____		
5. I'm proud of my school work._____		
6. I'm popular with kids my own age._____		
7. My parents usually consider my feelings._____		
8. I'm doing the best work that I can._____		
9. I can usually take care of myself._____		
10. My parents expect too much of me._____		
11. It's pretty tough to be me._____		
12. No one pays much attention to me at home._____		
13. I'm not doing as well in school as I'd like to._____		
14. I can make up my mind and stick to it._____		
15. I have a low opinion of myself._____		
16. I don't like to be with other people._____		
17. There are many times when I'd like to leave home._____		
18. I often feel upset in school._____		
19. I often feel ashamed of myself._____		
20. My teacher makes me feel I'm not good enough._____		
21. I get upset easily when I'm scolded._____		
22. I usually feel as if my parents are pushing me._____		
23. Things usually don't bother me._____		
24. I can't be depended on._____		

APPENDIX A.9

SUPPLEMENTAL QUESTIONS ASKED OF SUBJECTS IN ALL EXPERIMENTAL CONDITIONS

On page 210 occurs the Supplemental Questions asked of all subjects in the experimental conditions. The oral instructions are given in the text.

QUESTIONS

1. A few days ago, when you took the Sizes and Shapes Test, there were 20 problems where you answered A or B. Some people were just now informed what their scores were. Other people's tests were not graded. If you were told what your score was, write the number you got wrong on the line below. If you weren't informed what your score was, write the number of problems you think you got wrong out of 20 on the line below (make your best guess.)

_____ wrong.

2. How good are you at sizing up new situations? (Circle one below.)

Very Good	Pretty Good	Fair	Pretty Poor	Very Poor

3. How important do you think it is for a young girl or boy to be able to size up new situations? (Circle one below.)

Very Important	Pretty Important	Neither Important nor Unimportant	Pretty Unimportant	Very Unimportant

4. How well did the other boys in you class do on the Sizes and Shapes Test? (Circle one below.)

Very Well	Pretty Well	Fair	Pretty Poorly	Very Poorly

5. How good are the other boys in your class in sizing up new situations? (Circle one below.)

Very Good	Pretty Good	Fair	Pretty Poor	Very Poor

6. How satisfied are you with your ability to size up new situations? (Circle one below.)

Very Satisfied	Pretty Satisfied	Neither Satisfied nor Dissatisfied	Pretty Dissatisfied	Very Dissatisfied

7. Think of the boys in your class who are your good friends. Circle the statement below that best describes how you feel about these friends.

Like them a whole lot	Like them pretty much	Like them a little	Dislike them a little	Dislike them pretty much	Dislike them very much

8. Think of the boys in your class who are not your good friends. Circle the statement that best describes how you feel about these other boys.

Like them a whole lot	Like them pretty much	Like them a little	Dislike them a little	Dislike them pretty much	Dislike them very much

APPENDIX B.1

RATIONALE AND DERIVATION OF SPECIAL FORMULA USED TO ESTIMATE A SUBJECT'S TRUE CHANGE

A special formula was developed by the author to estimate the true change in an individual's attribute. This computation was based on suggestions made by Cronbach and Furby (in press). This appendix details the rationale and derivation of the formula.

An individual has two measures, X_o and Y_o, taken on an attribute. Raw change $(Y_o - X_o)$ is unsuitable as a measure of change because the reliability of this difference is extremely low. This is because observed scores, especially for those who are extreme, are filled with large error components.

We wish to make an estimate of true change, D'_t, that is as free of error, that is of regression effects, as possible. There are two ways we can do this: (1) We can calculate the residual gain. The gain is the difference between the obtained Y and the Y predicted from X on the basis of the X on Y regression line. Thus, $Y_o - Y'$, where Y' is predicted in the usual way from X_o. (2) We can correct for the unreliability of measurement in X_o. All scores above the mean probably have plus errors while those below have minus errors. We can predict true X, X_t, from X_o using the formula $x_t = r_{xx'} x_o$ where $r_{xx'}$ is the

reliablity coefficient, $r_{xx'}$. In a similar way we can predict Y_t. An estimate of the true difference is, then, $Y_t - X_t$.

In the present formula, both principles are used. We first regress X_o onto X_t to get an estimate of true X. Then we use this estimate and the regression line of X_t on Y_t to predict Y'_t from X_o. This is a prediction of what true Y will be on the basis of the first measurement, X_o. We also find Y_t, an estimate of true Y from Y_o. Here we observe what was obtained for Y. The residualized estimate of true gain D'_t is given by $Y_t - Y'_t$. This procedure adjusts or corrects for the unreliability of X, the unreliability of Y, and the lack of perfect association between X and Y. D'_t has a lower variance and higher reliability than does $Y_o - X_o$.

Let us begin by finding Y'_t. The general formula for predicting Y from X is:

$$Y' = r_{xy} \frac{\sigma_x}{\sigma_y} (X - \bar{X}) + \bar{Y}$$

Here, we wish to predict Y_t from X_t. Specifically:

$$Y'_t = r_{x_t y_t} \frac{\sigma_{x_t}}{\sigma_{y_t}} (X_t - \bar{X}_t) + \bar{Y}_t \qquad (1)$$

We must regress X_o onto Y_t to predict true X. In general,

$$X_t = r_{xx'} X_o + (1 - r_{xx'}) \bar{X}_o \qquad (2)$$

where $r_{yy'}$ is the reliability ($\propto$) of the instrument. Thus we regress all terms to the mean, reducing their variance.

We need to know $r_{x_t y_t}$ and the standard deviations, σ_{x_t} and σ_{y_t}. Since Y_t is a linear transformation on Y_o:

$$r_{x_t y_t} = r_{x_o y_o}$$

$$\sigma_{x_t} = r_{xx'}\sigma_{x_o}$$

$$\sigma_{y_t} = r_{yy'}\sigma_{y_o}$$

$$\bar{Y}_t = r_{yy'}\bar{Y}_o + (1-r_{yy'})\bar{Y}_o = \bar{Y}_o$$

$$\bar{X}_t = \bar{X}_o$$

Thus, in terms of observed scores, (1) becomes

$$Y'_t = r_{x_o y_o} \frac{r_{xx'}\sigma_{x_o}}{r_{yy'}\sigma_{y_o}} (r_{xx'}(X_o - \bar{X}_o)) + \bar{Y}_o$$

The equation for Y_t was given in (2). Thus

$$D'_t = Y_t - Y'_t$$

$$D'_t = r_{yy'}Y_o + (1-r_{yy'})\bar{Y}_o - \left(r_{x_o y_o} \frac{r_{xx'}\sigma_{x_o}}{r_{yy'}\sigma_{y_o}} (r_{yy'}(X_o - \bar{X}_o)) + \bar{Y}_o \right)$$

Collecting terms we have:

$$D'_t = r_{yy'}Y_o - \left(\frac{r_{x_o y_o}(r_{xx'})^2 \sigma_{x_o}}{r_{yy'}\sigma_{y_o}} \right) X_o - \left(\frac{r_{x_o y_o}(r_{xx'})^2 \sigma_{x_o}}{r_{yy'}\sigma_{y_o}} \right) \bar{X}_o - r_{yy'}\bar{Y}_o$$

Over all Ss, the mean of $D'_t = 0$. We wish it, however, to equal $\bar{Y}_t - \bar{X}_t$, which is to say $\bar{Y}_o - \bar{X}_o$. Since the sum of the random errors equals 0, $\bar{Y}_t - \bar{X}_t$ is a true estimate of mean change. Thus we add this, $\bar{Y}_o - \bar{X}_o$, to each subject's score; the result is:

$$D'_t = r_{yy'}Y_o - \left(\frac{r_{x_o y_o}\left(r_{xx'}\right)^2 \sigma_{x_o}}{r_{yy'}\sigma_{y_o}}\right) X_o$$

$$- \left(1 - \frac{r_{x_o y_o}\left(r_{xx'}\right)^2 \sigma_{x_o}}{r_{yy'}\sigma_{y_o}}\right) \bar{X}_o - (1 - r_{yy'})\bar{Y}_o$$

The constants which are substituted ($r_{xx'}$, $r_{x_o y_o}$, $r_{yy'}$, σ_{x_o}, σ_{y_o}, $\bar{X}_o$ and $\bar{Y}_o$) are those from the control group, or from the sample as a whole. The X_o and Y_o variables are from a given subject and they yield his estimated change score.